Confronting the HIV/AIDS pandemic with a Functional Cure: SELY, a System to End Lentivirus-mediated immunodeficiencY

Author: **Rojo**pitiavana Solofondrazaintsoanirina*
**Correspondence:* solofo.rojo@gmail.com
DOI: 10.5281/zenodo.10809028, *published online:* March 12, 2024

Table of Contents

Disclaimer

The information provided in this document is intended for educational and informational purposes only. The author and publisher of this document are not responsible for any consequences that may arise from the use, misuse, or application of the techniques, methods, or instructions described herein. Instructions and guidelines outlined in this document are based on the author's understanding and knowledge as of the publication date. Scientific knowledge and technologies are constantly evolving, and readers are advised to consult with appropriate experts and references before conducting any experiments described in this document. Mention of specific trademarks or registered trademarks in this document does not imply endorsement, affiliation, or association with their holders. Trademarks belong to their respective owners and are used for identification purposes only. Readers should exercise caution and judgment when applying the information provided in this document. The author and publisher do not assume any responsibility for errors, omissions, or inaccuracies that may be present in the content. Additionally, readers are advised to comply with all applicable laws, regulations, and ethical considerations related to scientific experiments, inventions, and intellectual property. By using this document, readers acknowledge that they do so at their own risk and agree to hold the author and publisher harmless from any liability, claims, or damages that may arise from the use of the information provided herein.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

Introduction

When reverse transcriptase was discovered in 1970 within the Rous Sarcoma Virus (RSV), it astonished the scientific community by revealing how a single-strand RNA sequence could transform into double-strand DNA through a biological process called reverse transcription [1,2]. Thirteen years later, further research on human oncogenic retroviruses led to the identification of the Human Immunodeficiency Virus (HIV), comprising two species of *Lentivirus*, namely HIV-1 and HIV-2, which infect humans and lead to Acquired ImmunoDeficiency Syndrome (AIDS) [3,4]. AIDS manifests as the depletion of the CD4+ T cell population and the onset of various diseases associated with a progressively failing and chronically activated immune system, such as opportunistic infections and cancer.

According to UNAIDS estimates, approximately 39 million people were living with HIV in 2022, with 1.3 million newly infected individuals and 630,000 AIDS-related deaths recorded [5]. Presently, HIV remains incurable, but it can be effectively managed with highly potent single-tablet [6] or injectable [7,8] antiretroviral regimens. However, despite reducing the plasma virus to undetectable levels and halting disease progression, combination antiretroviral therapy (ART) fails to provide a cure due to the establishment of long-lived and persistent HIV reservoirs, resulting in viral rebound upon cessation of ART treatment within weeks. The HIV genome undergoes rapid mutation during treatment, leading to the emergence of drug-resistant strains. Additionally, escape mutants may arise, evading neutralization even in vaccinated individuals or HIV-infected patients who develop broad neutralizing antibodies (bNAbs).

The field of HIV vaccine research and development still faces numerous clinical failures [9], and an effective prophylactic HIV vaccine has yet to receive regulatory approval. However, thanks to years of scientific progress and the introduction of new biotechnological tools, promising new HIV vaccine candidates [10–13] and other therapeutic approaches are actively being pursued to combat the HIV/AIDS epidemic. These include CAR-T cell therapy [14–16], peptides and proteins inhibitors [17–20], immunotherapy [21,22], gene therapy [15,23–26], and CRISPR-Cas-mediated viral DNA excision [15,27,28].

Instances of HIV treatments leading to a cure are rare. In all cases, curative therapy involves a conditioning regimen followed by transplantation with CCR5-Δ32 hematopoietic stem cells [16,29–32]. However, this treatment carries significant risks for the patient, is expensive, not widely scalable, and does not protect against CXCR4-tropic viruses [33]. To address the longstanding HIV/AIDS epidemic, I have developed SELY (**S**ystem to **E**nd **L**entivirus-mediated immunodeficiency). SELY not only tackles major issues with current HIV treatment approaches but also acts as a novel therapeutic tool, working alongside the immune system to neutralize the virus and functionally cure HIV-infected patients.

Principle of working

The SELY therapeutic system is a conditionally replicative, conditionally active HIV-1 or HIV-2-derived lentiviral vector comprising a therapeutic RNA with a payload comprising sequences encoding one or multiple anti-HIV peptides or polypeptides. Each sequence is separated by a self-cleaving 2A peptide site. The payload cannot be translated due to native translation start sequences and a translational blocker positioned before its own translation start sequence (see Table S1 and FIGURE 1 and FIGURE 2 (A) and (B)). Upon transduction, the therapeutic RNA is reverse-transcribed, the translational blocker is deleted with a probability P_{act} (probability of activation). The newly synthesized therapeutic DNA is subsequently integrated into the cell's nucleus.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

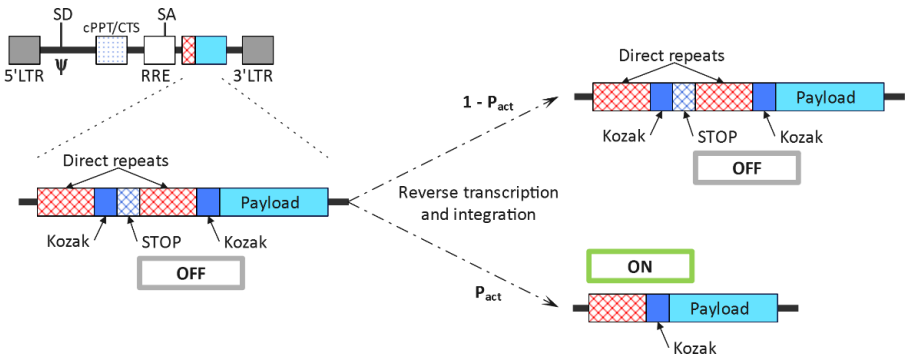

FIGURE 1: The key elements of an exemplary SELY therapeutic system include the translational blocker, which consists of direct repeats, Kozak sequences, and STOP signals. Reading is from 5' to 3'. The first Kozak sequence acts as the point where translation begins. When reverse transcription occurs, one of the direct repeats, the first Kozak, and the STOP signal are removed. This action leaves the remaining Kozak sequence as the new starting point for translation, activating the therapeutic payload (ON). If no deletions happen, the payload remains inactive (OFF). ψ: Packaging sequence; 5'LTR: 5 prime long terminal repeat; SD: Splice donor 1; cPPT/CTS: central polypurine tract/termination sequences; RRE: Rev-Response Element; SA: Splice acceptor 7; Kozak: Kozak consensus sequence; STOP: sequence comprising one or multiple stop codons (TAG, TGA, TAA); Payload: Sequence of the therapeutic payload; OFF: the payload is not translatable; ON: the payload is translatable; P_{act}: Probability of activation.

In cells where the translational blocker remains intact, the therapeutic DNA integrated into the cell is transcribed into therapeutic RNAs, but their therapeutic payload is unable to be translated. Subsequent infection by HIV leads to the mobilization of these therapeutic RNAs into infectious virions capable of transmitting to new hosts, sharing the same tropism as the infecting HIV strain (see FIGURE 2 (C)). This therapeutic system can also function as a Defective Interfering Particle (DIP), competing with HIV for available target cells and cellular resources [34,35].

However, in cells where the translational blocker is deleted, the therapeutic payload becomes active and is expressed. This renders the transduced cell resistant to HIV infections (FIGURE 2 (D)). The translational blocker comprises two identical sequences known as direct repeats, separated by an intervening sequence containing a translation start sequence and one or multiple in-frame stop codons (FIGURE 1). The translation start sequences are either Kozak consensus sequences or the 5' regions of HIV-1 open reading frames, serving as sites for translation initiation.

During reverse transcription, one of the direct repeat sequences and the intervening sequence are deleted with a probability referred to as P_{act}. This probability depends on the length of the direct repeats. For instance, direct repeats of 117, 284, and 971 nucleotides were deleted with probabilities of 0.068, 0.199, and 0.87, respectively [36–38]. Lowering the probability of activation (P_{act}) increases the proportion and number of transduced cells capable of serving as a therapeutic reservoirr.

Even if a cell contains a translationally active therapeutic payload, the transcribed therapeutic RNAs can still be mobilized if the cell becomes infected by HIV particles. This can occur due to inadequate expression of the therapeutic payload, a high viral load, or resistance developed by the infecting HIV strain.

A therapeutic reservoir refers to a cell containing one or multiple integrated SELY therapeutic DNAs with an intact translational blocker. Upon infection by HIV-1 or -2, the transcribed therapeutic RNAs compete with HIV genomic RNA for packaging into newly assembled virions. These virions then transduce new hosts, converting them into therapeutic reservoirs or HIV-resistant cells. Because HIV infection is necessary for SELY mobilization, a therapeutic reservoir may enter latency or become dormant, functioning as a sentinel cell that monitors and responds to any resurgence of HIV by producing SELY vectors.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

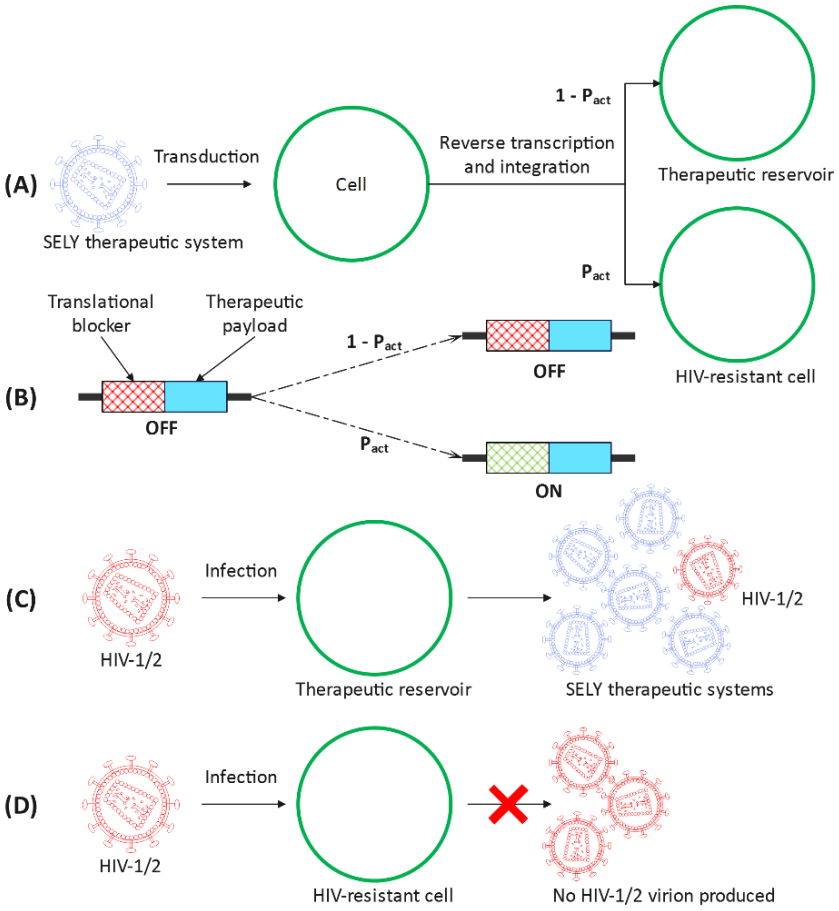

FIGURE 2: Working principle of SELY therapeutic system. (A) A cell that is transduced with the SELY therapeutic system can either become a therapeutic reservoir or resistant to HIV. (B) The translation of the therapeutic payload is initially blocked by the translational blocker. However, during the process of reverse transcription, there is a chance that the translational blocker may be deleted with a certain probability (P_{act}), allowing for the translation of the payload to occur. (C) When the therapeutic reservoir cell is infected by HIV, the SELY therapeutic system is mobilized into HIV virions. (D) An HIV-resistant cell is non-permissive for HIV replication.

Lentivectors can be pseudotyped with glycoproteins (GPs) from various enveloped viruses to adjust their ability to target specific cell types [39]. Two commonly used glycoproteins are VSV-G, known for its broad tropism and stability, and modified RD114, which facilitates efficient transduction of hematopoietic stem cells [40]. A SELY therapeutic system pseudotyped with a glycoprotein X forms what is called X-pseudotyped SELY vector, X SELY therapeutic system, or X-SELY vector.

The range of cells that can be transduced with SELY is extensive, including CD4+-CCR5+/CXCR4+ cells and CD34+ stem cells. Transducing CD34+ cells is especially appealing because once re-engrafted into a patient's bone marrow, they become self-renewable progenitor cells capable of evolving into therapeutic reservoirs and HIV-resistant immune cells (FIGURE 3). Cells, CD34+ hematopoietic progenitors, and lymphocyte T containing one or more genomically integrated or episomally maintained SELY therapeutic systems are referred to as SELY cells, SELY CD34+ cells, and SELY T cells, respectively.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

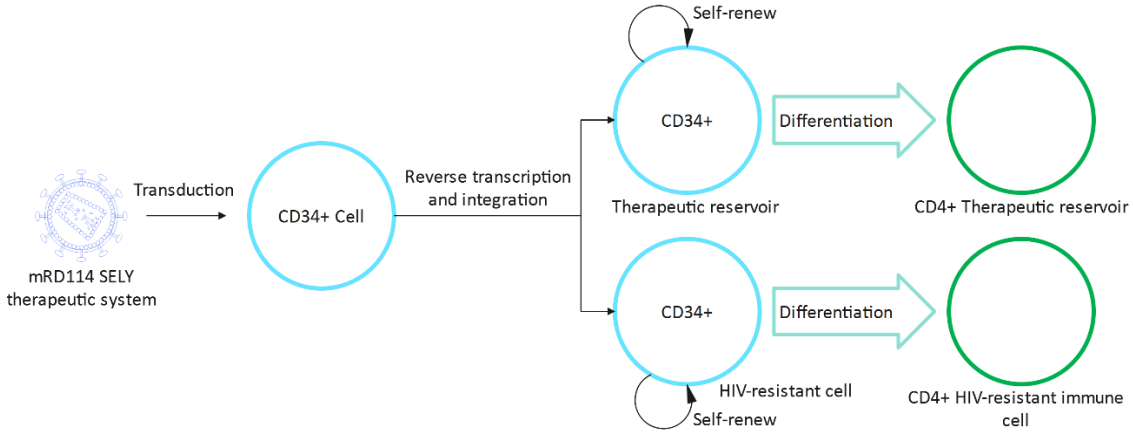

FIGURE 3: Modified RD114 (mRD114)-pseudotyped SELY therapeutic system (or mRD114-SELY vectors) can transduce CD34+ stem cells. For illustration purpose, the CD34+ cells can self-renew and differentiate into CD4+ lymphocyte cells that can be either therapeutic reservoir or HIV-resistant.

CD4+-CCR5+/CXCR4+ T cells are highly vulnerable to HIV infection. When stimulated by antigen-presenting cells, they undergo vigorous clonal expansion, resulting in the generation of numerous identical cells that are highly susceptible to both HIV [41] and SELY replication. Acting as a defective interfering particle, SELY can effectively compete with HIV, spreading more rapidly. This results in a large portion of CD4+ cell clones harboring integrated SELY therapeutic DNA. During the contraction phase, most of the remaining CD4+ cell clones are eliminated. The surviving clones that transition into memory cells can either become a therapeutic reservoir or develop resistance to HIV (FIGURE 4).

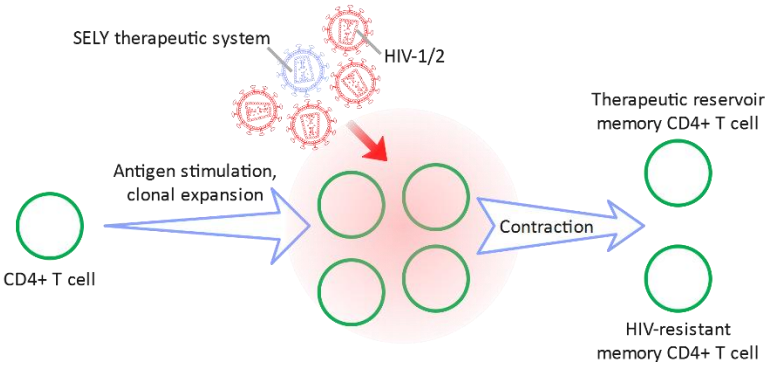

FIGURE 4: SELY therapeutic system can act as a Defective Interfering Particle (DIP) that competes against HIV during their replication in highly permissive clonally expanded CD4+ T cells. During the contraction phase, most of the CD4+ cell clones are eliminated and the remaining memory CD4+ T cells become either therapeutic reservoir or acquire resistance to HIV.

SELY vectors can be made integrase-defective, meaning the reverse-transcribed SELY therapeutic DNA exists as an episome inside the nucleus of the transduced cell. If the transduced cell is subsequently infected by HIV, the episome becomes actively transcribed. The newly produced SELY vectors have an intact integrase and share the same cellular tropism as the infecting HIV strain (FIGURE 5). Integrase-defective SELY vectors, when injected intravenously, are considered safer than their integrase-intact counterparts, particularly if they have a broad spectrum of cellular tropism. However, they do not persist for extended periods in cells undergoing mitosis.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

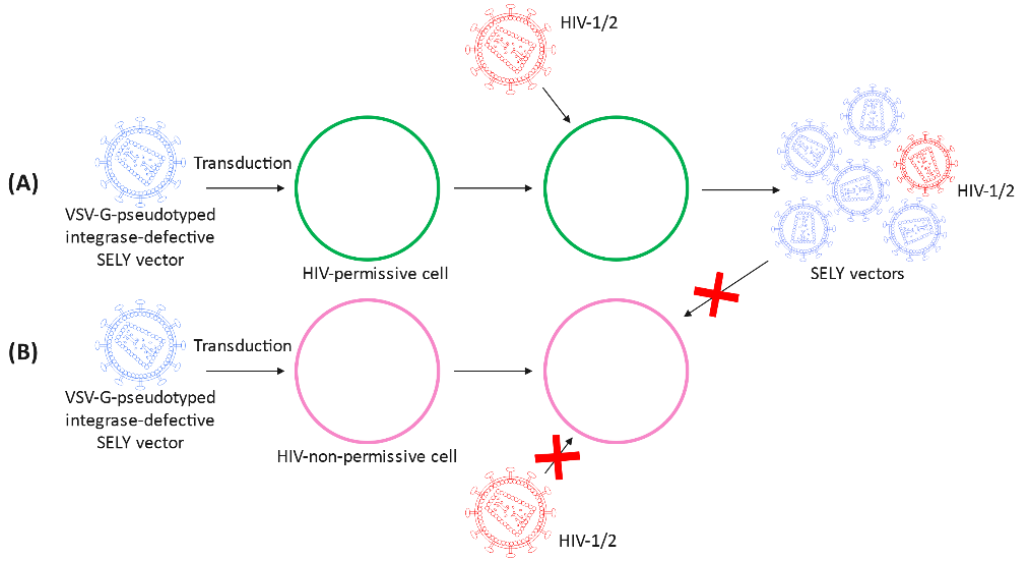

FIGURE 5: Illustration of VSV-G-pseudotyped integrase-defective SELY vectors transducing HIV-premissive and non-permissive cells. (A) When the transduced HIV-permissive cell is infected by HIV, the produced SELY vectors have an intact integrase and possess the same tropism as the infecting HIV strain. (B) The HIV-non-permissive cell (transduced or not) is refractory to infection by HIV as well as the newly produced SELY vectors.

Design

SELY therapeutic system

SELY therapeutic system is a lentivector (FIGURE 6) comprising, from 5' to 3':

(a) Promoter 1, a promoter that mediates the transcription of the therapeutic RNA;
(b) R, U5 (5' untranslated region), the packaging sequence, and a full or truncated Gag sequence of HIV-1 or -2;
(c) Cis-acting HIV sequences;
(d) The translational blocker and the payload;
(e) The U3 (3' untranslated region), R, and U5 sequences of HIV-1 or -2.

Promoter 1 is constitutive and includes CMV, PGK, HIV-1 U3, and RSV promoters. Its 3' region is truncated, so to initiate transcription at the 5' end of the 5' R sequence. Splice Donor 1 (SD1) is already present within the packaging sequence. The Gag sequence can either be full length or truncated, containing a translation start sequence and internal in-frame stop codons (FIGURE 6). Cis-acting sequences comprise a 118-nucleotide segment of HIV Pol containing the central polypurine tract/termination sequences, a portion of HIV Env containing the Rev-Response Element (RRE), and additional heterologous sequences that enhance viral titers and expression, such as the Woodchuck Hepatitis Virus (WHV) Posttranscriptional Regulatory Element (WPRE) [42]. Preferably, the 3' U3 region is either full-length or a hybrid sequence containing a heterologous promoter sequence or a segment of the long terminal repeat region from other *Retroviridae* family viruses [43]. Therapeutic lentiviral systems with intact U3 regions have already undergone clinical testing [44,45].

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

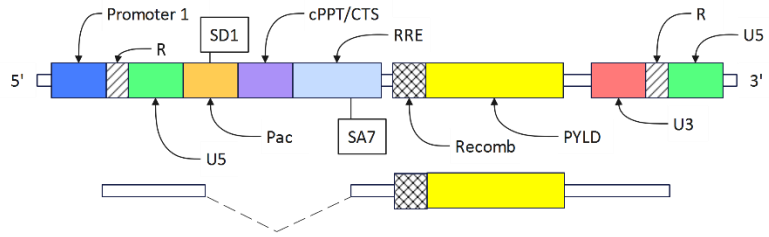

FIGURE 6: Components of an exemplary SELY therapeutic system. An SD1-SA7-spliced RNA sequence is shown below. Pac: sequence comprising the Packaging sequence; SD1: Splice Donor 1; cPPT/CTS: central polypurine tract/termination sequences; RRE: sequence comprising a Rev-Response Element; SA7: Splice Acceptor 7; Recomb: translational blocker or recombination site; PYLD: therapeutic payload sequence.

An exemplary translational blocker (FIGURE 7) is a sequence comprising, form 5' to 3':

(a) SeqA, a first direct repeat sequence;
(b) Kozak1, a translational start sequence or a Kozak consensus sequence;
(c) STP, a sequence comprising one or multiple stop codon (TGA, TAA, TAG) in-frame with the ATG sequence of kozak1;
(d) seqB, a second direct repeat sequence;
(e) Kozak2, a translational start sequence or a Kozak consensus sequence.

The percent identity between SeqA and SeqB ranges from 100% to less, and they have the same length. Following reverse-transcription, there's a chance that one of the direct repeats, kozak1, and STP are deleted, with a probability termed P_{act}. This probability can be adjusted by altering the lengths of SeqA-Kozak1, SeqB-Kozak2, or STP, and by changing the percent identity of SeqA-Kozak1 and SeqB-Kozak2..

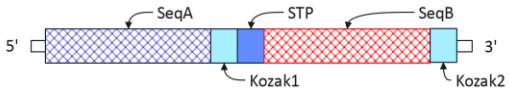

FIGURE 7: An exemplary translational blocker sequence.

After reverse-transcription and integration, the 3' U3 sequence is duplicated at the 5' end of the integrated therapeutic DNA, facilitating the transcription of therapeutic RNAs. These RNAs can undergo splicing at the SD1-SA7 site, removing the translation start site of the Gag sequence (FIGURE 6). If Kozak1 is deleted, Kozak2 becomes the new site for translation initiation.

The translational blocker may consist of one or multiple Kozak1 sequences, with the most 3' one followed by one or more stop codons. An incomplete translational blocker, lacking one of its direct repeats, is referred to as an OFF-translational blocker, a deleted translational blocker, or a translational blocker in the OFF state. Conversely, an intact translational blocker is called an ON-translational blocker, or an active translational blocker.

The payload comprises sequences encoding one or more therapeutic polypeptides, each potentially separated by a flexible linker or a 2A self-cleavable peptide (2A peptide or 2APEP) [46–48]. These payloads aren't limited to anti-HIV peptides or proteins; they can also include agonists that enhance HIV replication. Examples of agonists include HIV Tat proteins and targeted protein degradation (TPD) systems targeting APOBEC3G. Anti-APOBEC3G peptides and proteins help minimize SELY inactivation by hypermutation, while Tat polypeptides serve as potent latency-reversing agents [49,50] and strong transactivators, significantly increasing the production of therapeutic RNAs.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

The segment of the payload that we want to express constitutively can be duplicated and inserted at the 3' end of the first Kozak (FIGURE 8), followed by one or more in-frame stop codons. This segment always precedes any part of the payload we wish to express conditionally (FIGURE 8 (B)). Examples of sequences suitable for constitutive expression include those encoding Tat (FIGURE 9), drug-resistance, or surface marker (RQR8, tEGFR, tNGFR, CD20) polypeptides.

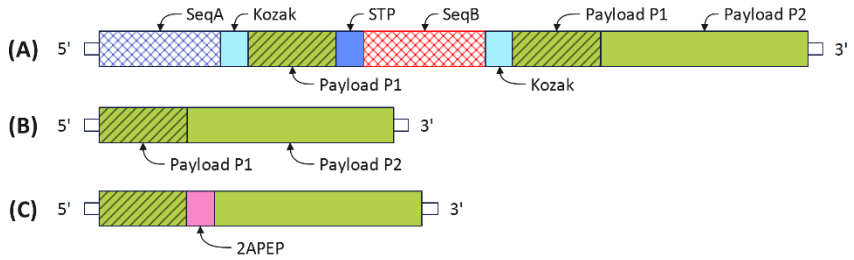

FIGURE 8: A segment of the payload can be constitutively expressed. (A) The segment (Payload P1 or 'Payload portion 1') we want to express constitutively is inserted 3' of the first Kozak. SeqA, Kozak, and Payload P1 make up the first direct repeat, while SeqB, Kozak, and Payload P1 form the second. Introducing silent mutations into Payload P1 of the second repeat can lower the value of P_{act}. (B) The payload's structure: Payload P1 is always 5' of Payload P2, the part we wish to express conditionally. (C) Payload P1 and P2 can be separated by a 2A self-cleaving peptide sequence.

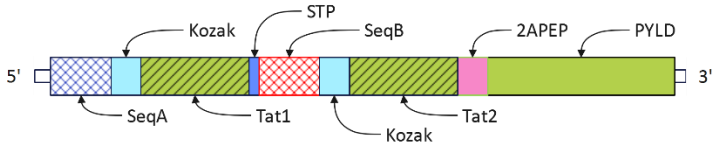

FIGURE 9: Tat polypeptide (encoded by Tat1) is constitutively expressed. A 2A self-cleaving peptide sequence (2APEP) separates Tat from the payload (PYLD). Tat1 and Tat2 encode the same Tat polypeptide, are of the same length, and can have 100% or less sequence identity.

The SELY therapeutic system may include a second payload (PYLD2) controlled by an Internal Ribosome Entry Site (IRES) sequence [51,52] (FIGURE 10). PYLD2 translation is unaffected by the presence of translation start sites, translational blockers, or upstream introns. PYLD2 can encode various polypeptides, including surface markers, drug-resistance proteins, safety switches (iCASP9 [53], DESNAses [54], RQR8 [55], Thymidine kinase [56]), fluorescent polypeptides, HIV latency reversing agents, viral accessory proteins (interferon antagonists, cell-cycle arresters, anti-apoptotic agents, etc.), prodrug-activating enzymes, or combinations thereof.

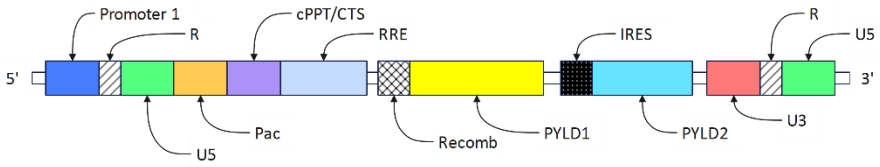

FIGURE 10: An exemplary SELY therapeutic system with a second payload (PYLD2) whose translation is independent of the first payload (PYLD1).

The SELY therapeutic system may also feature an internal promoter that independently transcribes the payload sequence, regardless of Tat presence (FIGURE 11). . The sequence containing the Rev-Response Element (RRE) can be inserted downstream of the payload (PYLD), generating transcripts that act as decoys for HIV Rev proteins [57].

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

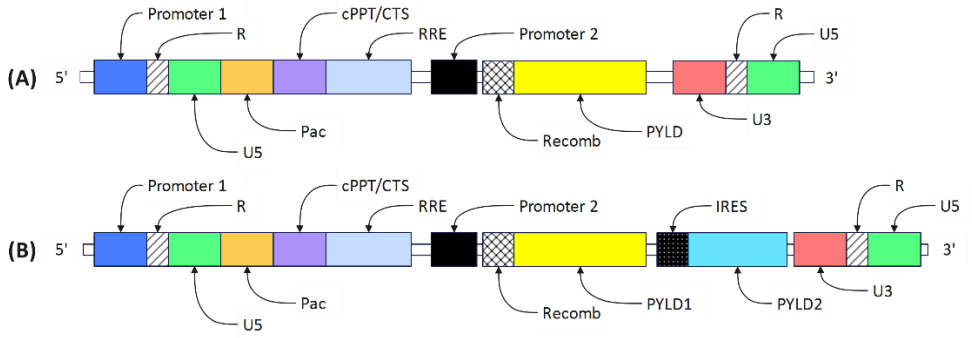

FIGURE 11: An exemplary SELY therapeutic system comprising an internal promoter (Promoter 2). (A) The internal promoter drives the transcription of Recomb and PYLD regardless of Tat presence. (B) The system further comprise a second Payload (PYLD2) under the translational control of an internal Ribosome Entry Site (IRES).

The Tat polypeptide plays a crucial role in viral replication as it significantly boosts the production of viral transcripts driven by the U3 promoter. It can be expressed independently from the payload (PYLD) by utilizing HIV's differential splicing ability (FIGURE 12). An HIV-native Tat sequence, including the upstream 'splice acceptor 3' (SA3) sequence, can be inserted downstream of the 118-nucleotide cPPT/CTS sequence from the HIV Pol sequence, which contains a branch point and a polypyrimidine tract essential for full intron definition.

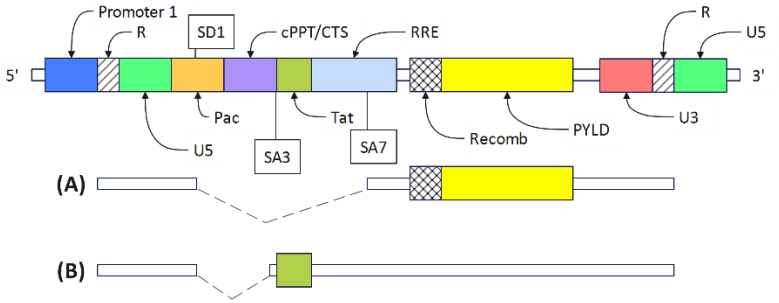

FIGURE 12: An exemplary SELY therapeutic system capable of constitutively expressing Tat. (A) Splicing at SD1-SA7 removes the translation start site of Gag and Tat, allowing the payload to be translated if the translational blocker is deleted. (B) Splicing at SD1-SA3 remove the translation start site of Gag, enabling Tat translation. SA3: Splice Acceptor 3; Tat: sequence encoding HIV Tat.

A second payload (PYLD2) can be co-expressed with Tat by inserting it downstream of the Tat sequence. Both sequences can be separated by a cis-separator (Cis sep), which can be a 2A self-cleaving peptide sequence or an Internal Ribosome Entry Site (IRES).

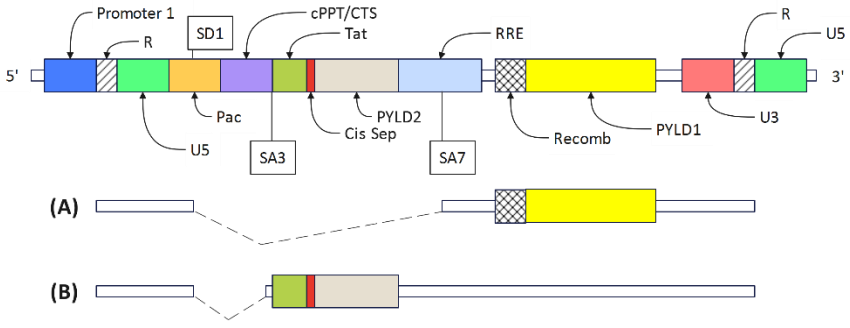

FIGURE 13: A second Payload (PYLD2) can be inserted 3' of Tat sequence. The two sequences can be separated by a Cis-separator (Cis sep), which can be a 2A self-cleaving peptide sequence or an Internal Ribosome Entry Site (IRES). (A) Splicing at SD1-SA7

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

removes the translation start site of Gag and Tat, allowing PYLD1 translation if the translational blocker is deleted. (B) Splicing at SD1-SA3 removes the translation start site of Gag, enabling the joint-expression of Tat and the second payload. SA3: Splice Acceptor 3; Tat: sequence encoding HIV Tat; PYLD2: Payload 2; Cis sep: Cis-separator.

The SELY therapeutic system may also include one or more artificial micro-RNA (miRNA) cassettes. These can be inserted into intronic regions, such as between the Gag sequence and cPPT/CTS, between cPPT/CTS and RRE, or downstream of the payload sequence and upstream of 3'-U3 (FIGURE 14). To mitigate the risk of deletion due to repeated sequences, miRNA cassettes with different backbone sequences can be combined to form a polycistronic miRNA cassette [58].

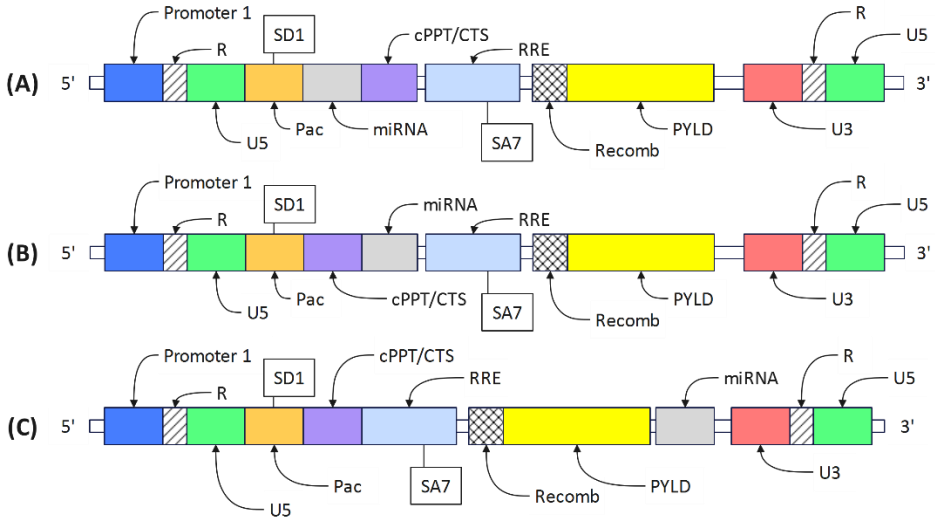

FIGURE 14: Exemplary SELY therapeutic system comprising one or multiple artificial micro-RNA (miRNA) cassettes. The miRNA cassette can be inserted between Gag and cPPT/CTS (A), cPPT/CTS and RRE (B), or downstream of the payload (C).

The SELY therapeutic system can be delivered and integrated into the cell's genome using a transposon (FIGURE 15). Examples of transposons include those derived from PiggyBac [59], Sleeping Beauty [60], Frog Prince [61], Tol2 [62,63], and PiggyBat [64]. SELY therapeutic sequence (an example is illustrated in FIGURE 16) is inserted between the transposon's 5' and 3' flanking ends, allowing for its mobilization and genomic integration in the presence of an appropriate transposase polypeptides. Examples of suitable transposon plasmids are: transposon Petit Plasmids [65,66], PiggyBac Nanoplasmid [67], and Tol2 transposon [63].

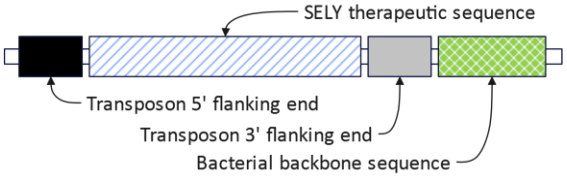

FIGURE 15: A transposon comprising a SELY therapeutic sequence. Transposon 5' and 3' flanking ends are sufficient and necessary for the transposase-mediated mobilization and genomic integration of SELY into the host's DNA. The 5' and 3' ends of this transposon sequence can be covalently linked to form a circular plasmid. The bacterial backbone sequence is required for maintenance and replication of the plasmid in an appropriate bacterial host.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

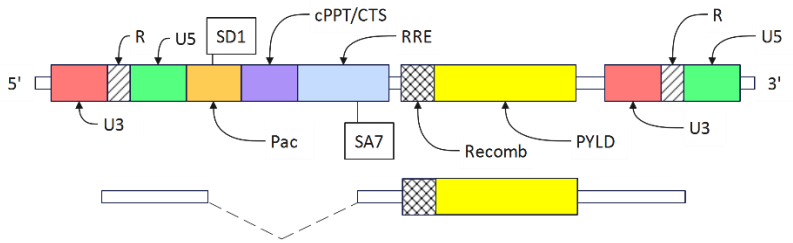

FIGURE 16: An exemplary SELY therapeutic sequence. Promoter1 is replaced with the HIV-1 U3 sequence. A polyadenylation sequence such as SV40 or BGH polyA is optional and can be inserted at the 3' end (after the 3' U5 region) of this sequence.

Translational blocker

One purpose of a translational blocker is to prevent translation initiation at a downstream (or 3') translation initiation site or Kozak consensus sequence. It can be built by selecting a sequence of desired length to serve as the first direct repeat. Following this, a translation start site or a Kozak consensus sequence is added at its 3' end, followed by one or more codons, then one or more in-frame stop codons (TGA, TAA, TAG), and finally a second direct repeat of the same length as the first direct repeat with a percent identity of 100% or less. Non-limiting examples of sequences suitable for direct repeats can be found between the translation start site (TSS) and including the start codon (ATG) of a gene or a truncation thereof.

Another purpose of a translational blocker is to halt translation at one of consecutive in-frame stop codons, ensuring the downstream coding sequence remains untranslated. This type of blocker can be constructed by selecting a sequence of desired length as the first direct repeat, followed by one or more in-frame stop codons (TGA, TAA, TAG), and then a second direct repeat of the same length and percent identity of 100% or less. Ideally, a translation initiation site or Kozak consensus sequence is positioned 5' of the first direct repeat. In the example below, the polypeptide sequence of a 72-aa HIV Tat and its corresponding *Homo sapiens* codon-optimized nucleotide sequence are shown. A 110-nt sequence is highlighted in green. The translational blocker comprises The Kozak consensus sequence (highlighted in cyan), the Tat coding sequence (in bold and red), and the two 110-nt direct repeats (highlighted in green). Here, translation is initiated at the ATG codon of the Kozak consensus sequence and is terminated at any one the four in-frame stop codons.

Tat72 polypeptide:
MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQRRRAHQNSQTHQASLSKQ

Tat72 nucleotide sequence:
ATGGAGCCCGTGGACCCTAGGCTGGAGCCCTGGAAACATCCCGGATCTCAACCCAAGACCGCCTGCACCAACTGCTACTGCAA
GAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAACGAAGGCAGAGG
AGGAGGGCCCACCAGAACAGCCAGACCCACCAAGCTAGCCTTAGCAAGCAG

Translational blocker sequence (recomb110_kozak_tat72):
GCCGCCAACCATGGAGCCCGTGGACCCTAGGCTGGAGCCCTGGAAACATCCCGGATCTCAACCCAAGACCGCCTGCACCAAC
TGCTACTGCAAGAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAAC
GAAGGCAGAGGAGGAGGGCCCACCAGAACAGCCAGACCCACCAAGCTAGCCTTAGCAAGCAGTGATAATAGTGA**TGTGCTT**
CATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAACGAAGGCAGAGGAGGAGGGCCCACCAGAACAGCCAGACCC
ACCAAGCTAGCCTTAGCAAGCAG

Translational blocker sequence (recomb110_kozak_tat72_gsg-p2a) with a self-cleaving 2A peptide sequence (highlighted in yellow) at its 3' end:

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

GCCGCCAACCATGGAGCCCGTGGACCCTAGGCTGGAGCCCTGGAAACATCCCGGATCTCAACCCAAGACCGCCTGCACCAAC
TGCTACTGCAAGAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAAC
GAAGGCAGAGGAGGAGGGCCCACCAGAACAGCCAGACCCACCAAGCTAGCCTTAGCAAGCAGTGATAATAGTGATGTGCTT
CATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAACGAAGGCAGAGGAGGAGGGCCCACCAGAACAGCCAGACCC
ACCAAGCTAGCCTTAGCAAGCAGGGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGAGATGTTGAGGAGAACCC
TGGACCT

The length of the Tat72 nucleotide sequence including the Kozak consensus sequence is 226-nt. To construct a translational blocker with a longer 276-nt direct repeat sequences, a 50-nt sequence is inserted at the 5' end of the Kozak consensus sequence (226 + 50 = 276). This entire sequence of length 276 is the first direct repeat.

First direct repeat:
ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCGCCGCCAACCATGGAGCCCGTGGACCCTAGGCTG
GAGCCCTGGAAACATCCCGGATCTCAACCCAAGACCGCCTGCACCAACTGCTACTGCAAGAAGTGCTGCTTCCACTGCCAGGT
GTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAACGAAGGCAGAGGAGGAGGGCCCACCAGAACAGCC
AGACCCACCAAGCTAGCCTTAGCAAGCAG

The translational blocker (recomb276_kozak_tat72_gsg-p2a) thus comprises the first direct repeat, the in-frame stop codons (doubly underlined), the second direct repeat (of the same length and percent identity as the first direct repeat), and accessorily a self-cleaving 2A peptide sequence (highlighted in yellow):
ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCGCCGCCAACCATGGAGCCCGTGGACCCTAGGCTG
GAGCCCTGGAAACATCCCGGATCTCAACCCAAGACCGCCTGCACCAACTGCTACTGCAAGAAGTGCTGCTTCCACTGCCAGGT
GTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAACGAAGGCAGAGGAGGAGGGCCCACCAGAACAGCC
AGACCCACCAAGCTAGCCTTAGCAAGCAGTGATAATAGTGAACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAAC
AGACACCGCCGCCAACCATGGAGCCCGTGGACCCTAGGCTGGAGCCCTGGAAACATCCCGGATCTCAACCCAAGACCGCCTG
CACCAACTGCTACTGCAAGAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGG
AAGAAACGAAGGCAGAGGAGGAGGGCCCACCAGAACAGCCAGACCCACCAAGCTAGCCTTAGCAAGCAGGGCAGCGGCG
CCACCAACTTCAGCCTGCTGAAGCAGGCCGGAGATGTTGAGGAGAACCCTGGACCT

In the case of recomb110_kozak_tat72_gsg-p2a and recomb276_kozak_tat72_gsg-p2a, the Tat72 sequence downstream of the first Kozak consensus sequence can be readily translated. The expressed Tat72 is a potent latency reversing agent, less toxic than the full-length Tat, and can mediate enhanced therapeutic RNA production [50,68].

The translational blocker can be combined with the RRE sequence. Here, Env_RRE, a portion of HIV-1 Env sequence comprising the RRE (highlighted in green) is shown below. Splice acceptor 7 (SA7) is highlighted in yellow. A sequence of length 112-nt can serve as direct repeat and is highlighted in cyan. A direct repeat longer than 112-nt can be obtained by adding one or more nucleotides after the 3' end of this sequence.

GATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCAT
TAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTG
GGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATA
GTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCT
CCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTT
GCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGG
ACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAA
TTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATG
ATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTA
TCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGA
GACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATC

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

Below is an example of a portion of HIV-1 Env sequence comprising the RRE wherein the two direct repeats—each one is 100-nt long—are highlighted in cyan, the Kozak consensus sequence is underlined (the start codon is in bold), splice acceptor 7 (SA7) is highlighted in yellow, and five in-frame stop codons are in red. We will refer to it as **Env_RRE_recomb100**. During reverse-transcription, one of the direct repeat is deleted, the stop codons are removed, and only the last Kozak consensus sequence is retained. To lower the probability of deletion, direct repeats of shorter length can be chosen. Alternatively, one or more nucleotides of the second direct repeat can be altered such that the first and second direct repeats have a percent identity of less than 100%. To increase the probability of deletion, the spacing between the two direct repeats can be increased. Alternatively, the direct repeats can be lengthened by inserting: 1) one or more nucleotides at the 5' end of the Kozak consensus sequences and/or; 2) one or more codons at the 3' end of the start codon of the Kozak consensus sequences.

GATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCAT
TAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTG
GGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATA
GTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCT
CCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTT
GCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGG
ACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAA
TTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATG
ATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTA
TCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGA
GACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGCCGCCACC**ATG**GCCAGATGATAATAGTGATAACCCGACAGGC
CCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGCC
GCCACC**ATG**

The length of a direct repeat is crucial as it determines the probability of activation (P_{act}). As presented previously, 117-, 284-, and 971-nt identical direct repeats were deleted with a probability of 0.068, 0.199, and 0.87 respectively [36–38]. In another study, a 1,333-, 788-, and 383-nt direct repeat are deleted with a probability of 0.93, 0.85, and 0.28 to 0.4, respectively [69]. The percent identity of the two direct repeats also influences P_{act}. A 156-nt direct repeat with a 100% homology (percent identity) is deleted with a probability of 0.149. A reduction in homology to 95, 91, 82, 73, 63, and 58% yields a deletion probability of 0.096, 0.039, 0.0077, <0.0004, <0.0005, and <0.0008, respectively [36]. SELY therapeutic system can have a P_{act} ranging from 0.001 to 0.99.

Payloads

The payload, located 3' of the translational blocker, consists of sequences encoding one or multiple polypeptides. Each polypeptide can be separated by a linker, such as flexible linkers (e.g., the sequence GGGGS repeated N times), self-cleaving 2A peptides, or protease cleavage sites. Translation of the payload occurs only when the translational blocker undergoes deletion of one of its direct repeats and the in-frame stop codons. The payload may encode various polypeptides, including anti-HIV proteins, HIV agonists or latency reversing agents (ex: Tat), surface marker polypeptides, drug-resistance proteins, drug-sensitizing proteins, safety switches, fluorescent polypeptides, viral accessory proteins, prodrug-activating enzymes, and any combination thereof. Polypeptides can be retained in the cytosol or secreted into the extracellular space by adding an appropriate signal peptide at their N-terminus.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

Implementation

An exemplary SELY therapeutic system (CONSTRUCT_001) consists of (from 5' to 3'): prom1_CMV; R_U5_Psi; tGag255; cPPT/CTS; Env_RRE_recomb50; PYLD; fU3_R_U5; L_SV40_polyA

Exemplary DNA and polypeptide sequences are in Supplementary materials. Env_RRE_recomb50 comprises two 50-nt direct repeat sequences (highlighted in cyan) as shown below.

GATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCAT
TAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTG
GGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATA
GTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCT
CCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTT
GCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGG
ACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAA
TTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATG
ATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTA
TCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGA
GACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGCCGCCACC**ATG**GCCAGATGATAATAGTGATAACAGATCCATTC
GATTAGTGAACGGATCTCGACGGTATCGCCGCCACC**ATG**

The payload (PYLD) comprises a reverse transcriptase inhibitor (RTI) [18]; a self-cleaving 2A peptide; HIV-1 Tat66; a self-cleaving 2A peptide; and a plasma membrane-anchored fusion inhibitor (mC46) [19]. Its sequence is as follow: RTI (KETWETWWTE); GSG-P2A (GSGATNFSLLKQAGDVEENPGP); Tat66 (MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQRRRAHQNSQTHQ); GSG-T2A (GSGEGRGSLLTCGDVEENPGP); LNGFR signal peptide (MGAGATGRAMDGPRLLLLLLLGVSLGGA); C46 peptide (WMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF); thIgG2 hinge (ERKCCVECPPCPAPPVAGP); CD34 TMD (LIALVTSGALLAVLGITGYFLMNRRSWSPTGERLELEP)

Payload polypeptide (251-aa):

MKETWETWWTEGSGATNFSLLKQAGDVEENPGPMEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKR
RQRRRAHQNSQTHQGSGEGRGSLLTCGDVEENPGPMGAGATGRAMDGPRLLLLLLLGVSLGGAWMEWDREINNYTSLIHSLIEES
QNQQEKNEQELLELDKWASLWNWFERKCCVECPPCPAPPVAGPLIALVTSGALLAVLGITGYFLMNRRSWSPTGERLELEP

Payload nucleotide sequence:

ATGAAGGAGACCTGGGAGACCTGGTGGACCGAGGGCAGCGGCGCCACAAATTTCAGCTTGCTGAAGCAGGCCGGAGACGTG
GAGGAGAACCCTGGCCCTATGGAACCTGTGGATCCTAGGCTGGAGCCCTGGAAGCACCCTGGCAGCCAACCCAAGACAGCTT
GTACCAACTGCTACTGCAAGAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGGA
AGAAGAGGAGGCAAAGGAGGAGGGCCCACCAGAACAGCCAGACCCACCAGGGATCTGGAGAAGGAAGAGGCAGCCTGCTG
ACCTGCGGAGACGTGGAGGAGAATCCTGGACCTATGGGCGCTGGAGCCACAGGAAGGGCTATGGATGGACCTCGGCTGCTGT
TGTTGCTGCTGCTGGGCGTGAGCTTAGGAGGAGCCTGGATGGAATGGGATAGGGAGATCAACAACTACACCAGCCTGATCCA
CAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTG
GAACTGGTTCGAGAGGAAGTGCTGCGTGGAGTGTCCTCCTTGTCCTGCTCCTCCTGTTGCTGGCCCGTTGATCGCCCTGGTTAC
AAGCGGAGCCCTGCTTGCCGTTCTTGGCATCACCGGCTACTTCCTGATGAACAGGAGGAGCTGGAGCCCTACCGGCGAGAGGC
TGGAGCTTGAGCCT

The sequence of CONSTRUCT_001 is thus (CMV promoter is highlighted in yellow, direct repeats are highlighted in cyan, payload highlighted in green, in-frame stop codons are doubly underlined, the late SV40 polyadenylation signal is in bold and green):

GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATA
ACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAAC

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

GCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT
GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCC
TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCG
GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA
AAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGCGCGTTTTG
CCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAAT
AAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTC
AGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAGGACTC
GGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGA
AGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAG
GGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTA
GAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATA
TAATACAGTAGCAACCCTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAA
CAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAATTCAAAATTTTCGGGTTTATTGATCTTCAGACCTGGAGGAGGA
GATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGC
AAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAG
CACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGC
TGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTG
GAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAAT
GCTAGTTGGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACA
AGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGC
AAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTT
AAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACC
CCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAAC
GGATCTCGACGGTATCGCCGCCACCATGGCCAGATGATAATAGTGATAACAGATCCATTCGATTAGTGAACGGATCTCGACGG
TATCGCCGCCACCATGAAGGAGACCTGGGAGACCTGGTGGACCGAGGGCAGCGGCGCCACAAATTTCAGCTTGCTGAAGCAG
GCCGGAGACGTGGAGGAGAACCCTGGCCCTATGGAACCTGTGGATCCTAGGCTGGAGCCCTGGAAGCACCCTGGCAGCCAAC
CCAAGACAGCTTGTACCAACTGCTACTGCAAGAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCA
GCTACGGCAGGAAGAAGAGGAGGCAAAGGAGGAGGGCCCACCAGAACAGCCAGACCCACCAGGGATCTGGAGAAGGAAGA
GGCAGCCTGCTGACCTGCGGAGACGTGGAGGAGAATCCTGGACCTATGGGCGCTGGAGCCACAGGAAGGGCTATGGATGGA
CCTCGGCTGCTGTTGTTGCTGCTGCTGGGCGTGAGCTTAGGAGGAGCCTGGATGGAATGGGATAGGGAGATCAACAACTACA
CCAGCCTGATCCACAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
GGGCCAGCCTGTGGAACTGGTTCGAGAGGAAGTGCTGCGTGGAGTGTCCTCCTTGTCCTGCTCCTCCTGTTGCTGGCCCGTTGA
TCGCCCTGGTTACAAGCGGAGCCCTGCTTGCCGTTCTTGGCATCACCGGCTACTTCCTGATGAACAGGAGGAGCTGGAGCCCTA
CCGGCGAGAGGCTGGAGCTTGAGCCTTGATAAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTT
TTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACA
AGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGT
ACCAGTTGAGCCAGATAAGATAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATG
GATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGG
AGTACTTCAAGAACTGCTGACATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGG
ACTGGGGAGTGGCGAGCCCTCAGATCCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTG
AGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGC
CCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAG**CAGACATGATAA**
GATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGC
TTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGG
TGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTA

CONSTRUCT_001 is 3,128-nt long without the CMV promoter and SV40 polyadenylation sequence. The full CONSTRUCT_001 can be cloned into an empty vector for maintenance and replication in an appropriate *E. coli* host. A desired amount of plasmids can be obtained using commercially available plasmid purification

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

kits. The plasmids can serve in many downstream applications such as the production of high titer lentivirus vector pseudotypes.

Another exemplary SELY therapeutic system (CONSTRUCT_002) consists of (from 5' to 3'): prom1_CMV; R_U5_Psi; tGag255; cPPT/CTS; Env_RRE; recomb150_kozak_tat66_gsg-p2a; PYLD; fU3_R_U5; L_SV40_polyA.

Tat66 consists of the first 66-aa of HIV-1 Tat and its nucleotide sequence is (a 150-nt sequence serving as the first direct repeat is highlighted in cyan):
ATGGAGCCCGTGGACCCTAGGCTGGAGCCCTGGAAACATCCCGGATCTCAACCCAAGACCGCCTGCACCAACTGCTACTGCAA
GAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAACGAAGGCAGAGG
AGGAGGGCCCACCAGAACAGCCAGACCCACCAA

The sequence of recomb150_kozak_tat66_gsg-p2a consists of a Kozak consensus sequence followed by Tat66 (it comprises the first 150-nt direct repeat), three in-frame stop codons, a 150-nt direct repeat, and a self-cleaving 2A peptide (GSG-P2A):
GCCGCCACCATGGAGCCCGTGGACCCTAGGCTGGAGCCCTGGAAACATCCCGGATCTCAACCCAAGACCGCCTGCACCAACTG
CTACTGCAAGAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAACGAA
GGCAGAGGAGGAGGGCCCACCAGAACAGCCAGACCCACCAATGATGATAACAACCCAAGACCGCCTGCACCAACTGCTACTG
CAAGAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGGAAGAAACGAAGGCAG
AGGAGGAGGGCCCACCAGAACAGCCAGACCCACCAAGGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGAGAT
GTTGAGGAGAACCCTGGACCT

The payload (PYLD) comprises a reverse transcriptase and integrase inhibitor (RT&INI); a self-cleaving 2A peptide; and a plasma membrane-anchored fusion inhibitor (mC46). Its sequence is as follow: RT&INI (VEAIIRILQQLLFIH); GSG-T2A (GSGEGRGSLLTCGDVEENPGP); LNGFR signal peptide (MGAGATGRAMDGPRLLLLLLLGVSLGGA); C46 peptide (WMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF); thIgG2 hinge (ERKCCVECPPCPAPPVAGP); CD34 TMD (LIALVTSGALLAVLGITGYFLMNRRSWSPTGERLELEP)

Payload polypeptide:
VEAIIRILQQLLFIHGSGEGRGSLLTCGDVEENPGPMGAGATGRAMDGPRLLLLLLLGVSLGGAWMEWDREINNYTSLIHSLIEESQN
QQEKNEQELLELDKWASLWNWFERKCCVECPPCPAPPVAGPLIALVTSGALLAVLGITGYFLMNRRSWSPTGERLELEP

Payload nucleotide sequence:
GTGGAGGCCATCATCAGAATCCTGCAGCAGCTGCTGTTCATCCACGGCAGCGGCGAGGGAAGAGGCTCTTTGCTGACCTGCGG
AGATGTTGAGGAGAACCCTGGACCTATGGGAGCTGGAGCCACAGGAAGGGCTATGGACGGACCTAGGCTGCTGCTTCTGCTG
CTGCTGGGAGTGAGCCTTGGAGGAGCCTGGATGGAGTGGGATAGGGAGATCAACAACTACACCAGCCTGATCCACAGCCTGA
TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGT
TCGAGAGGAAGTGCTGCGTGGAGTGCCCTCCCTGTCCTGCTCCTCCTGTTGCTGGACCTTTGATCGCCCTGGTGACAAGCGGCG
CCCTTCTGGCCGTTCTTGGCATTACAGGCTACTTCCTGATGAACAGGAGGAGCTGGAGCCCTACCGGCGAGAGGCTGGAGCTT
GAGCCT

The sequence of CONSTRUCT_002 is thus (CMV promoter is highlighted in yellow, Kozak consensus sequence is underlined, direct repeats are highlighted in cyan, payload highlighted in green, in-frame stop codons are doubly underlined, the late SV40 polyadenylation signal is in bold and green):
GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATA
ACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAAC
GCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT
GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCC

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCG
GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA
AAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGCGCGTTTTG
CCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAAT
AAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTC
AGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAGGACTC
GGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGA
AGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAG
GGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTA
GAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATA
TAATACAGTAGCAACCCTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAA
CAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAATTCAAAATTTTCGGGTTTATTGATCTTCAGACCTGGAGGAGGA
GATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGC
AAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAG
CACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGC
TGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTG
GAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAAT
GCTAGTTGGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACA
AGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGC
AAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTT
AAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACC
CCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAAC
GGATCTCGACGGTATCGCCGCCACC**ATG**GAGCCCGTGGACCCTAGGCTGGAGCCCTGGAAACATCCCGGATCTCAACCCAAGA
CCGCCTGCACCAACTGCTACTGCAAGAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACG
GCAGGAAGAAACGAAGGCAGAGGAGGAGGGCCCACCAGAACAGCCAGACCCACCAATGATGATAACAACCCAAGACCGCCT
GCACCAACTGCTACTGCAAGAAGTGCTGCTTCCACTGCCAGGTGTGCTTCATCACCAAGGCCCTGGGCATCAGCTACGGCAGG
AAGAAACGAAGGCAGAGGAGGAGGGCCCACCAGAACAGCCAGACCCACCAAGGCAGCGGCGCCACCAACTTCAGCCTGCTG
AAGCAGGCCGGAGATGTTGAGGAGAACCCTGGACCTGTGGAGGCCATCATCAGAATCCTGCAGCAGCTGCTGTTCATCCACGG
CAGCGGCGAGGGAAGAGGCTCTTTGCTGACCTGCGGAGATGTTGAGGAGAACCCTGGACCTATGGGAGCTGGAGCCACAGG
AAGGGCTATGGACGGACCTAGGCTGCTGCTTCTGCTGCTGCTGGGAGTGAGCCTTGGAGGAGCCTGGATGGAGTGGGATAGG
GAGATCAACAACTACACCAGCCTGATCCACAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGC
TGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGAGAGGAAGTGCTGCGTGGAGTGCCCTCCCTGTCCTGCTCCTCCT
GTTGCTGGACCTTTGATCGCCCTGGTGACAAGCGGCGCCCTTCTGGCCGTTCTTGGCATTACAGGCTACTTCCTGATGAACAGG
AGGAGCTGGAGCCCTACCGGCGAGAGGCTGGAGCTTGAGCCTTGATAAGGTACCTTTAAGACCAATGACTTACAAGGCAGCT
GTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCT
GTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGG
ATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTG
TGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGC
CCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGA
GGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCT
CTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGT
GCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCT
AGCAG**CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGA**
AATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTT
TCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTA

CONSTRUCT_002 is 3,228-nt long without the CMV promoter and SV40 polyadenylation sequence. Tat66 is constitutively expressed and should enhance the production of therapeutic RNAs in cells transduced with this version of SELY.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

CONSTRUCT_003 is an exemplary SELY therapeutic system consisting, from 5’ to 3’, of: prom1_CMV; R_U5_Psi; tGag255; cPPT/CTS; SA3_tat198$_{ex1}$; GSG-P2A_RQR8; Env_RRE_recomb150; PYLD; IRES; iCASP9; fU3_R_U5; L_SV40_polyA.

Env_RRE_recomb150 comprises two 150-nt direct repeats and has the sequence (a 26-nt sequence (in bold and orange) is inserted between the 3’ end of Env_RRE and the 5’ end of the first Kozak consensus sequence so the length of the direct repeats are 150-nt each):
GATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCAT
TAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTG
GGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATA
GTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCT
CCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTT
GCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGG
ACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAA
TTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATG
ATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTA
TCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGA
GACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCCTTCTGACACAACTGTGTTCACTAGCGCCGCCACC**ATG**GCCAGA
TGATAATAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGA
GACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCCTTCTGACACAACTGTGTTCACTAGCGCCGCCACC**ATG**

RQR8 is a cell surface marker polypeptide [55] with the sequence (signal peptide is underlined):
GTSLLCWMALCLLGADHADACPYSNPSLCSGGGGSELPTQGTFSNVSTNVSPAKPTTTACPYSNPSLCSGGGGSPAPRPPTPAPTIAS
QPLSLRPEACRPAAGGAVHTRGLDFACDIYIWAPLAGTCGVLLLSLVITLYCNHRNRRRVCKCPRPVV

GSG-P2A_RQR8 has the sequence:
GSGATNFSLLKQAGDVEENPGPGTSLLCWMALCLLGADHADACPYSNPSLCSGGGGSELPTQGTFSNVSTNVSPAKPTTTACPYSN
PSLCSGGGGSPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACDIYIWAPLAGTCGVLLLSLVITLYCNHRNRRRVCKCPR
PVV

GSG-P2A_RQR8 nucleotide sequence:
GGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGAGATGTTGAGGAGAACCCTGGCCCTGGCACAAGCCTGCTGT
GCTGGATGGCCCTTTGCTTGTTGGGAGCCGACCACGCCGATGCCTGTCCTTACAGCAACCCTAGCCTGTGCTCTGGCGGAGGA
GGCAGCGAGCTTCCTACACAGGGAACCTTCAGCAACGTGAGCACCAACGTGAGCCCTGCCAAGCCCACCACCACCGCTTGTCCT
TACAGCAACCCTAGCCTGTGCAGCGGAGGAGGAGGATCTCCTGCTCCTAGGCCTCCTACACCTGCTCCTACCATCGCCAGCCAG
CCTCTGAGCCTTAGGCCTGAAGCTTGTAGGCCTGCTGCTGGCGGAGCTGTGCACACAAGAGGCCTGGATTTCGCCTGTGACAT
CTACATCTGGGCTCCCTTGGCCGGCACCTGCGGAGTTCTGCTGCTGTCTCTTGTGATCACCCTGTACTGCAACCACAGGAACAG
GAGGAGGGTGTGCAAGTGCCCTAGGCCCGTGGTG

iCASP9 is the inducible human caspase9 safety switch [53] with the sequence:
GVQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKVDSSRDRNKPFKFMLGKQEVIRGWEEGVAQMSVGQRAKLTISPDYAYGAT
GHPGIIPPHATLVFDVELLKLESGGGSGVDGFGDVGALESLRGNADLAYILSMEPCGHCLIINNVNFCRESGLRTRTGSNIDCEKLRRRF
SSLHFMVEVKGDLTAKKMVLALLELARQDHGALDCCVVVILSHGCQASHLQFPGAVYGTDGCPVSVEKIVNIFNGTSCPSLGGKPKL
FFIQACGGEQKDHGFEVASTSPEDESPGSNPEPDATPFQEGLRTFDQLDAISSLPTPSDIFVSYSTFPGFVSWRDPKSGSWYVETLDDI
FEQWAHSEDLQSLLLRVANAVSVKGIYKQMPGCFNFLRKKLFFKTS

iCAPS9 nucleotide sequence:
GGCGTGCAGGTGGAGACCATCAGCCCTGGCGACGGCAGGACATTCCCTAAGAGAGGCCAGACCTGCGTTGTGCACTACACCG
GCATGCTGGAGGACGGCAAGAAGGTGGACAGCAGCAGGGACAGGAACAAGCCCTTCAAGTTCATGCTGGGCAAGCAGGAGG
TGATCAGAGGCTGGGAGGAGGGCGTTGCTCAGATGAGCGTGGGACAAAGGGCCAAGCTGACCATCAGCCCTGACTACGCCTA
CGGCGCTACAGGACATCCCGGCATCATCCCTCCTCACGCCACATTGGTGTTCGACGTGGAGCTGCTGAAGCTGGAGAGCGGAG

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

GAGGATCTGGCGTGGATGGCTTCGGAGACGTTGGAGCCTTGGAGAGCCTGAGAGGCAACGCCGACCTGGCTTACATCCTGAG
CATGGAGCCCTGTGGCCACTGCCTGATCATCAACAACGTGAACTTCTGCAGGGAGAGCGGCCTGAGGACCAGGACCGGCAGC
AACATCGATTGTGAGAAGCTGAGGAGGAGGTTCAGCAGCCTGCACTTCATGGTGGAGGTGAAGGGCGACCTGACCGCCAAGA
AGATGGTGCTGGCCTTGCTGGAGCTTGCTAGGCAGGACCACGGCGCTCTTGACTGCTGTGTGGTGGTGATCCTGAGCCACGGC
TGTCAGGCCAGCCACCTTCAGTTCCCTGGAGCTGTGTACGGAACCGACGGCTGTCCCGTGAGCGTGGAGAAGATCGTGAACAT
CTTCAACGGCACCAGCTGCCCTAGCCTGGGCGGCAAGCCCAAGCTGTTCTTCATTCAGGCCTGCGGAGGCGAACAGAAGGACC
ACGGCTTCGAGGTGGCTTCTACCAGCCCTGAGGACGAGTCTCCTGGCAGCAACCCTGAGCCTGACGCTACACCTTTCCAGGAG
GGCCTTAGGACCTTCGACCAGCTGGACGCCATCAGCTCTCTGCCCACACCCAGCGATATCTTCGTGAGCTACAGCACCTTCCCTG
GCTTCGTGAGCTGGAGGGACCCTAAGAGCGGCTCTTGGTACGTGGAGACCCTGGACGACATCTTCGAGCAGTGGGCCCACAG
CGAGGATCTGCAGAGCCTGCTGTTAAGGGTGGCCAACGCTGTGAGCGTGAAGGGCATCTACAAGCAGATGCCCGGCTGCTTC
AACTTCCTGAGGAAGAAGCTGTTCTTCAAGACCAGC

The IRES is taken from EMCV and has the sequence:
GTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGT
CTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACA
ACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGA
TACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCG
TATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACA
TGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAAT**ATG**

The payload (PYLD) comprises a reverse transcriptase and integrase inhibitor (RT&INI) [20]; a self-cleaving 2A peptide; and a plasma membrane-anchored fusion inhibitor (mC46). Its sequence is as follow: RT&INI (VEAIIRILQQLLFIH); GSG-T2A (GSGEGRGSLLTCGDVEENPGP); LNGFR signal peptide (MGAGATGRAMDGPRLLLLLLLGVSLGGA); C46 peptide (WMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF); thIgG2 hinge (ERKCCVECPPCPAPPVAGP); CD34 TMD (LIALVTSGALLAVLGITGYFLMNRRSWSPTGERLELEP)

Payload polypeptide:
MVEAIIRILQQLLFIHGSGEGRGSLLTCGDVEENPGPMGAGATGRAMDGPRLLLLLLGVSLGGAWMEWDREINNYTSLIHSLIEES
QNQQEKNEQELLELDKWASLWNWFERKCCVECPPCPAPPVAGPLIALVTSGALLAVLGITGYFLMNRRSWSPTGERLELEP

Payload nucleotide sequence:
GTGGAGGCCATCATCAGAATCCTGCAGCAGCTGCTGTTCATCCACGGCAGCGGCGAGGGAAGAGGCTCTTTGCTGACCTGCGG
AGATGTTGAGGAGAACCCTGGACCTATGGGAGCTGGAGCCACAGGAAGGGCTATGGACGGACCTAGGCTGCTGCTTCTGCTG
CTGCTGGGAGTGAGCCTTGGAGGAGCCTGGATGGAGTGGGATAGGGAGATCAACAACTACACCAGCCTGATCCACAGCCTGA
TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGT
TCGAGAGGAAGTGCTGCGTGGAGTGCCCTCCCTGTCCTGCTCCTCCTGTTGCTGGACCTTTGATCGCCCTGGTGACAAGCGGCG
CCCTTCTGGCCGTTCTTGGCATTACAGGCTACTTCCTGATGAACAGGAGGAGCTGGAGCCCTACCGGCGAGAGGCTGGAGCTT
GAGCCT

The sequence of CONSTRUCT_003 is thus (CMV promoter is highlighted in yellow, direct repeats are highlighted in cyan, payload highlighted in blue, in-frame stop codons are doubly underlined, the late SV40 polyadenylation signal is in bold and green):
GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATA
ACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAAC
GCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT
GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCC
TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCG
GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

AAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGCGCGTTTTG
CCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAAT
AAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTC
AGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAGGACTC
GGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGA
AGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAG
GGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTA
GAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATA
TAATACAGTAGCAACCCTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAA
CAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAATTCAAAATTTTCGGGTTTATTA**CAG**AATTGGGTGTCGACATAG
CAGAATAGGCGTTACTCGACAGAGGAGAGCAAGAA**ATG**GAGCCAGTAGATCCTAGACTAGAGCCCTGGAAGCATCCAGGAA
GTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGTTTCATAACAAAAGCCTTAG
GCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCATCAGAACAGTCAGACTCATCAAGGCAGCGGCGCC
ACCAACTTCAGCCTGCTGAAGCAGGCCGGAGATGTTGAGGAGAACCCTGGCCCTGGCACAAGCCTGCTGTGCTGGATGGCC
CTTTGCTTGTTGGGAGCCGACCACGCCGATGCCTGTCCTTACAGCAACCCTAGCCTGTGCTCTGGCGGAGGAGGCAGCGAGC
TTCCTACACAGGGAACCTTCAGCAACGTGAGCACCAACGTGAGCCCTGCCAAGCCCACCACCACCGCTTGTCCTTACAGCAAC
CCTAGCCTGTGCAGCGGAGGAGGAGGATCTCCTGCTCCTAGGCCTCCTACACCTGCTCCTACCATCGCCAGCCAGCCTCTGAG
CCTTAGGCCTGAAGCTTGTAGGCCTGCTGCTGGCGGAGCTGTGCACACAAGAGGCCTGGATTTCGCCTGTGACATCTACATCT
GGGCTCCCTTGGCCGGCACCTGCGGAGTTCTGCTGCTGTCTCTTGTGATCACCCTGTACTGCAACCACAGGAACAGGAGGAG
GGTGTGCAAGTGCCCTAGGCCCGTGGTG**TGATAA**GATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTG
AATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGA
AAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACG
CTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCA
TCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGC
TCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGG
AACAGATTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAG
AATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATA
ACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTA
TAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTT**CAG**ACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAA
GGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCCTTCTGACA
CAACTGTGTTCACTAGC**GCCGCCACC**ATGGCCAGATGATAATAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAA
GGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCCTTCTGACA
CAACTGTGTTCACTAGC**GCCGCCACC**ATGGTGGAGGCCATCATCAGAATCCTGCAGCAGCTGCTGTTCATCCACGGCAGCGGC
GAGGGAAGAGGCTCTTTGCTGACCTGCGGAGATGTTGAGGAGAACCCTGGACCTATGGGAGCTGGAGCCACAGGAAGGGCT
ATGGACGGACCTAGGCTGCTGCTTCTGCTGCTGCTGGGAGTGAGCCTTGGAGGAGCCTGGATGGAGTGGGATAGGGAGATCA
ACAACTACACCAGCCTGATCCACAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCT
GGACAAGTGGGCCAGCCTGTGGAACTGGTTCGAGAGGAAGTGCTGCGTGGAGTGCCCTCCCTGTCCTGCTCCTCCTGTTGCTG
GACCTTTGATCGCCCTGGTGACAAGCGGCGCCCTTCTGGCCGTTCTTGGCATTACAGGCTACTTCCTGATGAACAGGAGGAGCT
GGAGCCCTACCGGCGAGAGGCTGGAGCTTGAGCCT**TGATAA**GTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCC
CGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTC
GTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACC
TGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGA
GTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCAT
TGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCCCCGAAC
CACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAAT**ATG**GGCGTGCAGGTGGAGACCATCAGCCCTGGCGACGGCAGG
ACATTCCCTAAGAGAGGCCAGACCTGCGTTGTGCACTACACCGGCATGCTGGAGGACGGCAAGAAGGTGGACAGCAGCAGGG
ACAGGAACAAGCCCTTCAAGTTCATGCTGGGCAAGCAGGAGGTGATCAGAGGCTGGGAGGAGGGCGTTGCTCAGATGAGCG
TGGGACAAAGGGCCAAGCTGACCATCAGCCCTGACTACGCCTACGGCGCTACAGGACATCCCGGCATCATCCCTCCTCACGCCA
CATTGGTGTTCGACGTGGAGCTGCTGAAGCTGGAGAGCGGAGGAGGATCTGGCGTGGATGGCTTCGGAGACGTTGGAGCCTT
GGAGAGCCTGAGAGGCAACGCCGACCTGGCTTACATCCTGAGCATGGAGCCCTGTGGCCACTGCCTGATCATCAACAACGTGA

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

ACTTCTGCAGGGAGAGCGGCCTGAGGACCAGGACCGGCAGCAACATCGATTGTGAGAAGCTGAGGAGGAGGTTCAGCAGCC
TGCACTTCATGGTGGAGGTGAAGGGCGACCTGACCGCCAAGAAGATGGTGCTGGCCTTGCTGGAGCTTGCTAGGCAGGACCA
CGGCGCTCTTGACTGCTGTGTGGTGGTGATCCTGAGCCACGGCTGTCAGGCCAGCCACCTTCAGTTCCCTGGAGCTGTGTACGG
AACCGACGGCTGTCCCGTGAGCGTGGAGAAGATCGTGAACATCTTCAACGGCACCAGCTGCCCTAGCCTGGGCGGCAAGCCCA
AGCTGTTCTTCATTCAGGCCTGCGGAGGCGAACAGAAGGACCACGGCTTCGAGGTGGCTTCTACCAGCCCTGAGGACGAGTCT
CCTGGCAGCAACCCTGAGCCTGACGCTACACCTTTCCAGGAGGGCCTTAGGACCTTCGACCAGCTGGACGCCATCAGCTCTCTG
CCCACACCCAGCGATATCTTCGTGAGCTACAGCACCTTCCCTGGCTTCGTGAGCTGGAGGGACCCTAAGAGCGGCTCTTGGTAC
GTGGAGACCCTGGACGACATCTTCGAGCAGTGGGCCCACAGCGAGGATCTGCAGAGCCTGCTGTTAAGGGTGGCCAACGCTG
TGAGCGTGAAGGGCATCTACAAGCAGATGCCCGGCTGCTTCAACTTCCTGAGGAAGAAGCTGTTCTTCAAGACCAGCTGATAA
GGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAA
TTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACC
AGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCC
AATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGA
GGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCT
ACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCAT
ATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACC
CACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATC
CCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAG**CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACT**
AGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACA
AGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTC
TACAAATGTGGTA

CONSTRUCT_003 is 5,716-nt long without the CMV promoter and SV40 polyadenylation sequence. Tat66 [50] can be constitutively expressed after mRNA splicing at SD1-SA3. It should enhance the production of therapeutic RNAs in cells transduced with this version of SELY. RQR8 is co-expressed with Tat66 and serve as a selectable marker and safety switch. iCASP9 is also constitutively expressed and can induce apoptosis in the presence of small molecule ligands such as AP1903. When the translational blocker is OFF, the payload can be expressed after mRNA splicing at SD1-SA7.

CONSTRUCT_004 is an exemplary SELY therapeutic system similar in structure to CONSTRUCT_003 but having its CMV sequence (prom1_CMV) replaced with HIV-1 U3 sequence (prom1_U3). It consists, from 5' to 3', of: prom1_U3; R_U5_Psi; tGag255; cPPT/CTS; SA3_tat198$_{ex1}$; GSG-P2A_RQR8; Env_RRE_recomb150; PYLD; IRES; iCASP9; fU3_R_U5; L_SV40_polyA.

CONSTRUCT_004 can be inserted into a Piggybac transposon vector comprising an origin of replication and a selectable marker necessary for plasmid's replication and maintenance, and the positive selection of *E. coli* transformed with the vector. Alternatively, CONSTRUCT_004 can be flanked at both of its end with a transposon 5' and 3' flanking ends (Table S4) and then cloned into an empty vector.

CONSTRUCT_004_T consists of the sequence of CONSTRUCT_004 inserted between two PiggyBac transposon's 5' and 3' flanking ends: ePB_5_prime; prom1_U3; R_U5_Psi; tGag255; cPPT/CTS; SA3_tat198$_{ex1}$; GSG-P2A_RQR8; Env_RRE_recomb150; PYLD; IRES; iCASP9; fU3_R_U5; **L_SV40_polyA**; ePB_3_prime.

CONSTRUCT_004_T is 6,751-nt and has the sequence:
CGCAGCTAGATTAACCCTAGAAAGATAGTCTGCGTAAAATTGACGCATGCATTCTTGAAATATTGCTCTCTCTTTCTAAATAGCG
CGAATCCGTCGCTGTGCATTTAGGACATCTCAGTCGCCGCTTGGAGCTCCCGTGAGGCGTGCTTGTCAATGCGGTAAGTGTCAC
TGATTTTGAACTATAACAACCGCGTGAGTCAAAATGACGCATGATTATCTTTTACGTGACTTTTAAGATTTAACTCATACGATAA
TTATATTGTTATTTCGTGTTCTACTTACGTGATAACTTATTATATATATATTTTCTTGTTATAGATATCCTTCTGGAAGGGCTAATT
CACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCA

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

GGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCCA
ATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAG
GTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCTA
CAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCATA
TAAGCAGCTGCTTTTTGCCTGTACT**GGG**TCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCC
ACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCC
CTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTC
TCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGA
CTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAA
ATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAG
TTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAA
GAACTTAGATCATTATATAATACAGTAGCAACCCTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAAT
AGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAATTCAAAATTTTCGGGTTTATTA**CAG**A
ATTGGGTGTCGACATAGCAGAATAGGCGTTACTCGACAGAGGAGAGCAAGAA**ATG**GAGCCAGTAGATCCTAGACTAGAGCC
CTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGTT
TCATAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCATCAGAACAGTCAGACTC
ATCAAGGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGAGATGTTGAGGAGAACCCTGGCCCTGGCACAAGCC
TGCTGTGCTGGATGGCCCTTTGCTTGTTGGGAGCCGACCACGCCGATGCCTGTCCTTACAGCAACCCTAGCCTGTGCTCTGGC
GGAGGAGGCAGCGAGCTTCCTACACAGGGAACCTTCAGCAACGTGAGCACCAACGTGAGCCCTGCCAAGCCCACCACCACC
GCTTGTCCTTACAGCAACCCTAGCCTGTGCAGCGGAGGAGGAGGATCTCCTGCTCCTAGGCCTCCTACACCTGCTCCTACCATC
GCCAGCCAGCCTCTGAGCCTTAGGCCTGAAGCTTGTAGGCCTGCTGCTGGCGGAGCTGTGCACACAAGAGGCCTGGATTTCG
CCTGTGACATCTACATCTGGGCTCCCTTGGCCGGCACCTGCGGAGTTCTGCTGCTGTCTCTTGTGATCACCCTGTACTGCAACC
ACAGGAACAGGAGGAGGGTGTGCAAGTGCCCTAGGCCCGTGGTGTGATAAGATCTTCAGACCTGGAGGAGGAGATATGAG
GGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGA
AGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATG
GGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGC
TATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGAT
ACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTT
GGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAA
TACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTG
TGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATA
GTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTT**CAG**ACCCACCTCCCAACCCCGAGGG
GACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTC
GACGGTATCCTTCTGACACAACTGTGTTCACTAGCGCCGCCACC**ATG**GCCAGATGATAATAGACCCACCTCCCAACCCCGAGG
GGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCT
CGACGGTATCCTTCTGACACAACTGTGTTCACTAGCGCCGCCACC**ATG**GTGGAGGCCATCATCAGAATCCTGCAGCAGCTGCT
GTTCATCCACGGCAGCGGCGAGGGAAGAGGCTCTTTGCTGACCTGCGGAGATGTTGAGGAGAACCCTGGACCTATGGGAGCT
GGAGCCACAGGAAGGGCTATGGACGGACCTAGGCTGCTGCTTCTGCTGCTGCTGGGAGTGAGCCTTGGAGGAGCCTGGATGG
AGTGGGATAGGGAGATCAACAACTACACCAGCCTGATCCACAGCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACG
AGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTCGAGAGGAAGTGCTGCGTGGAGTGCCCTCCCTG
TCCTGCTCCTCCTGTTGCTGGACCTTTGATCGCCCTGGTGACAAGCGGCGCCCTTCTGGCCGTTCTTGGCATTACAGGCTACTTC
CTGATGAACAGGAGGAGCTGGAGCCCTACCGGCGAGAGGCTGGAGCTTGAGCCTTGATAAGTTATTTTCCACCATATTGCCGT
CTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAA
TGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCA
GGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACA
ACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGG
ATGCCCAGAAGGTACCCCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAA
AACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAAT**ATG**GGCGTGCAGGTGGAGACCAT
CAGCCCTGGCGACGGCAGGACATTCCCTAAGAGAGGCCAGACCTGCGTTGTGCACTACACCGGCATGCTGGAGGACGGCAAG
AAGGTGGACAGCAGCAGGGACAGGAACAAGCCCTTCAAGTTCATGCTGGGCAAGCAGGAGGTGATCAGAGGCTGGGAGGAG

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

GGCGTTGCTCAGATGAGCGTGGGACAAAGGGCCAAGCTGACCATCAGCCCTGACTACGCCTACGGCGCTACAGGACATCCCG
GCATCATCCCTCCTCACGCCACATTGGTGTTCGACGTGGAGCTGCTGAAGCTGGAGAGCGGAGGAGGATCTGGCGTGGATGG
CTTCGGAGACGTTGGAGCCTTGGAGAGCCTGAGAGGCAACGCCGACCTGGCTTACATCCTGAGCATGGAGCCCTGTGGCCACT
GCCTGATCATCAACAACGTGAACTTCTGCAGGGAGAGCGGCCTGAGGACCAGGACCGGCAGCAACATCGATTGTGAGAAGCT
GAGGAGGAGGTTCAGCAGCCTGCACTTCATGGTGGAGGTGAAGGGCGACCTGACCGCCAAGAAGATGGTGCTGGCCTTGCTG
GAGCTTGCTAGGCAGGACCACGGCGCTCTTGACTGCTGTGTGGTGGTGATCCTGAGCCACGGCTGTCAGGCCAGCCACCTTCA
GTTCCCTGGAGCTGTGTACGGAACCGACGGCTGTCCCGTGAGCGTGGAGAAGATCGTGAACATCTTCAACGGCACCAGCTGCC
CTAGCCTGGGCGGCAAGCCCAAGCTGTTCTTCATTCAGGCCTGCGGAGGCGAACAGAAGGACCACGGCTTCGAGGTGGCTTCT
ACCAGCCCTGAGGACGAGTCTCCTGGCAGCAACCCTGAGCCTGACGCTACACCTTTCCAGGAGGGCCTTAGGACCTTCGACCA
GCTGGACGCCATCAGCTCTCTGCCCACACCCAGCGATATCTTCGTGAGCTACAGCACCTTCCCTGGCTTCGTGAGCTGGAGGGA
CCCTAAGAGCGGCTCTTGGTACGTGGAGACCCTGGACGACATCTTCGAGCAGTGGGCCCACAGCGAGGATCTGCAGAGCCTGC
TGTTAAGGGTGGCCAACGCTGTGAGCGTGAAGGGCATCTACAAGCAGATGCCCGGCTGCTTCAACTTCCTGAGGAAGAAGCTG
TTCTTCAAGACCAGCTGATAAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAA
GGGGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCC
CTGATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAG
CCAGATAAGATAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGG
AGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAG
AACTGCTGACATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGT
GGCGAGCCCTCAGATCCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAG
CTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGT
GTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAG**CAGACATGATAAGATACATTGAT**
GAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAC
CATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTT
TTTTAAAGCAAGTAAAACCTCTACAAATGTGGTACGATAAAAGTTTTGTTACTTTATAGAAGAAATTTTGAGTTTTTGTTTTTTT
TTAATAAATAAATAAACATAAATAAATTGTTTGTTGAATTTATTATTAGTATGTAAGTGTAAATATAATAAAACTTAATATCTATT
CAAATTAATAAATAAACCTCGATATACAGACCGATAAAACACATGCGTCAATTTTACGCATGATTATCTTTAACGTACGTCACAA
TATGATTATCTTTCTAGGGTTAATCTAGCTGC

Purified plasmids bearing CONSTRUCT_004_T can be co-transfected with PiggyBac transposase expression plasmids into any desired mammalian cell types, including CD34+ hematopoietic progenitor cells, CD4+ T cells, or induced pluripotent stem cells.

CONSTRUCT_004_T_empty consists of the sequence of CONSTRUCT_004_T wherein the translational blocker and payload are removed and replaced with restriction sites (AfeI and PacI). Its sequence is as follow:

CGCAGCTAGATTAACCCTAGAAAGATAGTCTGCGTAAAATTGACGCATGCATTCTTGAAATATTGCTCTCTCTTTCTAAATAGCG
CGAATCCGTCGCTGTGCATTTAGGACATCTCAGTCGCCGCTTGGAGCTCCCGTGAGGCGTGCTTGTCAATGCGGTAAGTGTCAC
TGATTTTGAACTATAACAACCGCGTGAGTCAAAATGACGCATGATTATCTTTTACGTGACTTTTAAGATTTAACTCATACGATAA
TTATATTGTTATTTCGTGTTCTACTTACGTGATAACTTATTATATATATATTTTCTTGTTATAGATATCCTTCTGGAAGGGCTAATT
CACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCA
GGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCCA
ATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAG
GTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCTA
CAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCATA
TAAGCAGCTGCTTTTTGCCTGTACT**GGG**TCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCC
ACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCC
CTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTC
TCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGA
CTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAA
ATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAG

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

TTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAA
GAACTTAGATCATTATATAATACAGTAGCAACCCTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAAT
AGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAATTCAAAATTTTCGGGTTTATTA**CAG**A
ATTGGGTGTCGACATAGCAGAATAGGCGTTACTCGACAGAGGAGAGCAAGAA**ATG**GAGCCAGTAGATCCTAGACTAGAGCC
CTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGTT
TCATAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCATCAGAACAGTCAGACTC
ATCAAGGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGAGATGTTGAGGAGAACCCTGGCCCTGGCACAAGCC
TGCTGTGCTGGATGGCCCTTTGCTTGTTGGGAGCCGACCACGCCGATGCCTGTCCTTACAGCAACCCTAGCCTGTGCTCTGGC
GGAGGAGGCAGCGAGCTTCCTACACAGGGAACCTTCAGCAACGTGAGCACCAACGTGAGCCCTGCCAAGCCCACCACCACC
GCTTGTCCTTACAGCAACCCTAGCCTGTGCAGCGGAGGAGGAGGATCTCCTGCTCCTAGGCCTCCTACACCTGCTCCTACCATC
GCCAGCCAGCCTCTGAGCCTTAGGCCTGAAGCTTGTAGGCCTGCTGCTGGCGGAGCTGTGCACACAAGAGGCCTGGATTTCG
CCTGTGACATCTACATCTGGGCTCCCTTGGCCGGCACCTGCGGAGTTCTGCTGCTGTCTCTTGTGATCACCCTGTACTGCAACC
ACAGGAACAGGAGGAGGGTGTGCAAGTGCCCTAGGCCCGTGGTGTGATAAGATCTTCAGACCTGGAGGAGGAGATATGAG
GGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGA
AGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATG
GGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGC
TATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGAT
ACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTT
GGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAA
TACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTG
TGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATA
GTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTT**CAG**ACCCACCTCCCAACCCCGAGGG
GACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTC
GACGGTATCAGCGCTXXXXXXXXTTAATTAAGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGG
CCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGC
AGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGT
GCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTT
GTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATC
TGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGGCCCCCCGAACCACGGGGACG
TGGTTTTCCTTTGAAAAACACGATGATAAT**ATG**GGCGTGCAGGTGGAGACCATCAGCCCTGGCGACGGCAGGACATTCCCTAA
GAGAGGCCAGACCTGCGTTGTGCACTACACCGGCATGCTGGAGGACGGCAAGAAGGTGGACAGCAGCAGGGACAGGAACAA
GCCCTTCAAGTTCATGCTGGGCAAGCAGGAGGTGATCAGAGGCTGGGAGGAGGGCGTTGCTCAGATGAGCGTGGGACAAAG
GGCCAAGCTGACCATCAGCCCTGACTACGCCTACGGCGCTACAGGACATCCCGGCATCATCCCTCCTCACGCCACATTGGTGTT
CGACGTGGAGCTGCTGAAGCTGGAGAGCGGAGGAGGATCTGGCGTGGATGGCTTCGGAGACGTTGGAGCCTTGGAGAGCCT
GAGAGGCAACGCCGACCTGGCTTACATCCTGAGCATGGAGCCCTGTGGCCACTGCCTGATCATCAACAACGTGAACTTCTGCA
GGGAGAGCGGCCTGAGGACCAGGACCGGCAGCAACATCGATTGTGAGAAGCTGAGGAGGAGGTTCAGCAGCCTGCACTTCA
TGGTGGAGGTGAAGGGCGACCTGACCGCCAAGAAGATGGTGCTGGCCTTGCTGGAGCTTGCTAGGCAGGACCACGGCGCTCT
TGACTGCTGTGTGGTGGTGATCCTGAGCCACGGCTGTCAGGCCAGCCACCTTCAGTTCCCTGGAGCTGTGTACGGAACCGACG
GCTGTCCCGTGAGCGTGGAGAAGATCGTGAACATCTTCAACGGCACCAGCTGCCCTAGCCTGGGCGGCAAGCCCAAGCTGTTC
TTCATTCAGGCCTGCGGAGGCGAACAGAAGGACCACGGCTTCGAGGTGGCTTCTACCAGCCCTGAGGACGAGTCTCCTGGCAG
CAACCCTGAGCCTGACGCTACACCTTTCCAGGAGGGCCTTAGGACCTTCGACCAGCTGGACGCCATCAGCTCTCTGCCCACACC
CAGCGATATCTTCGTGAGCTACAGCACCTTCCCTGGCTTCGTGAGCTGGAGGGACCCTAAGAGCGGCTCTTGGTACGTGGAGA
CCCTGGACGACATCTTCGAGCAGTGGGCCCACAGCGAGGATCTGCAGAGCCTGCTGTTAAGGGTGGCCAACGCTGTGAGCGT
GAAGGGCATCTACAAGCAGATGCCCGGCTGCTTCAACTTCCTGAGGAAGAAGCTGTTCTTCAAGACCAGCTGATAAGGTACCTT
TAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCC
CAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCA
GGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCCAATAAAG
GAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGA
CAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCTACAAGG
GACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCATATAAGC

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

AGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGC
TTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAG
ACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAG**CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATG**
CAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAA
CAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAA
TGTGGTACGATAAAAGTTTTGTTACTTTATAGAAGAAATTTTGAGTTTTTGTTTTTTTTTAATAAATAAATAAACATAAATAAATT
GTTTGTTGAATTTATTATTAGTATGTAAGTGTAAATATAATAAAACTTAATATCTATTCAAATTAATAAATAAACCTCGATATACA
GACCGATAAAACACATGCGTCAATTTTACGCATGATTATCTTTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAATCTAGC
TGC

The XXXXXXXX between the two restriction sites represents any intervening sequence of any length with a minimal size of 4-nt. For example, the intervening sequence can comprise a sequence encoding the first 59 residues of *E. coli* β-galactosidase [70] under the control of a constitutive prokaryotic promoter. The intervening sequence is removed and replaced with a payload (the insert) during cloning. Plasmids are transfected into *E. coli* suitable for blue/white screening. When cultured on agar plates, bacterial colonies harboring the right vector and insert are white while those containing the intact vector are blue.

A sequence comprising a translational blocker and a payload can be cloned into a plasmid harboring CONSTRUCT_004_T_empty by applying traditional molecular cloning techniques and using PacI and AfeI as restriction enzymes. The payload sequence can be terminated with at least one in-frame stop codon. If PacI and AfeI restriction sites are inappropriate, they can be replaced with any unique restriction site(s) or a Multiple Cloning Site (MCS) sequence.

Production

Exemplary DNA sequences comprising SELY therapeutic sequence are disclosed above. A DNA sequence of interest can be split into N subsequence of a desired size, gene-synthesized, and then assembled using well known techniques like Gibson assembly [71], Golden Gate assembly [72], and In-Fusion cloning [73].

SELY therapeutic sequences can be cloned between the 5’ and 3’ flanking regions (transposon direct repeats) of a transposon vector, replicated in a bacterial host, and purified for downstream applications. Alternatively, SELY therapeutic sequences can be cloned into empty vectors containing replication origins and selectable markers. Recombinant vectors produced after bacterial transformation and plasmid purification can be used for SELY vector production. Any lentiviral vectors production protocols described in the art are appropriate for use herein, with the condition that the co-transfection protocol involves the use of a plasmid harboring SELY therapeutic sequence, an envelope plasmid encoding a desired glycoprotein (VSV-G, LCMV-G [74], mRD114-G [40], cocal-G [75]), and packaging plasmids encoding Gag-Pol and Rev.

Alternatively, SELY vectors can be produced by transfecting plasmids harboring SELY therapeutic sequence into packaging cells that express integrated Gag-Pol, Env (envelop) and Rev genes upon the introduction of a small molecule ligand (example: doxycycline [76], cumate [77]) into the culture medium.

Integrase-defective packaging plasmids [78] can be utilized to produce integrase-defective SELY vectors. These vectors, when reverse-transcribed, remain as non-integrated episomes in the transduced cell nucleus, ensuring safety. In cells infected with HIV, the episomes are actively transcribed, and the newly produced SELY vectors have a functional integrase and are CD4+-CCR5/CXCR4 tropic. In cells non-permissive for HIV infection, the episomal SELY therapeutic DNA is gradually eliminated after several rounds of mitosis.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

Uses

Administration by injection

Pharmaceutical compositions containing integrase-intact or integrase-defective SELY vectors can be administered intradermally, intravenously, or intrabonally to patients. Intradermal gel injections allow slow vector release and passive diffusion across lymphoid structures. Intravenous injection can lead to the transduction of cells of the kidneys, spleen, and liver, while intrabone injections deliver vectors to bone marrow compartments, transducing hematopoietic progenitor cells.

Hematopoietic stem cells can be transduced by mobilizing them into circulation with mobilization stimuli [79], followed by intravenous injection of pharmaceutical compositions containing SELY vectors. These vectors may encode a drug-resistance gene that confers survival advantage to transduced cells. For example, SELY vectors comprising a sequence encoding MGMT-P140K [80], DHFR-L22Y [81], and miRNA/shRNA targeting HPRT [82] can protect transduced cells against the lethal effect of drugs such as O6-benzylguanine, methotrexate, and 6-thioguanine, respectively.

Ex vivo transfection or transduction of cells

The aim of transduction or transfection is to integrate SELY therapeutic DNA into the desired cells, such as induced pluripotent stem cells (iPSCs), CD34+ hematopoietic progenitor cells, or CD4+ T cells. Transduction consists of contacting cells of interest with an appropriate amount of SELY vectors. Any transduction protocols described in the art, involving lentivectors, and involving mammalian cells, are appropriate for use herein, replacing the protocol's lentivector and cells with SELY vectord and the desired cells, respectively. Transfection consists of contacting cells of interests with an appropriate volume or amount of a transfection cocktail comprising an adequate amount and ratio of transposon plasmids bearing a SELY therapeutic sequence and the appropriate transposase expression vectors. Any transfection protocols described in the art, involving transposon plasmids, and involving mammalian cells, are appropriate for use herein, replacing the protocol's transposon vectors and cells with transposon vectors harboring a SELY therapeutic sequence and the desired cells, respectively.

Following transduction with SELY vectors or transfection with transposon vectors harboring SELY therapeutic sequences, the modified cells can be injected into patients or cultured to expand their numbers. For instance, transduced autologous CD4+ T cells or autologous SELY CD4+ T cells can be stimulated with anti-CD28, anti-CD3 and cultured for days in the presence of IL-2/IL-15 and optionally IL-7 to increase their initial number. The clonally expanded cells are then re-infused back into patients either once or through a series of injections. In a similar approach, autologous CD4+ T cells were transduced with VRX496 lentiviral vectors, clonally expanded, and then re-infused back into patients at a dose of 5×10^{9} to 1×10^{10} cells administered biweekly [45]. The cells were able to persist for up to 5 years in some patients.

In the following example, transformed autologous CD34+ hematopoietic progenitor cells or autologous SELY CD34+ cells are re-infused back into patients without requiring myeloablative, lymphodepleting, or any form of preconditioning regimens. In a similar scenario, retrovirally-transduced CD34+ hematopoietic progenitor cells were intravenously administered into patients, absent any preconditioning regimens. It resulted in the gene marking of approximately 0.01-0.38% of the circulating peripheral blood mononuclear cells [83].

Cell samples from HIV-infected patients may contain infectious HIV, which can be mitigated by culturing cells with HIV protease inhibitors. During transfection for SELY therapeutic DNA integration, experimental or approved anti-HIV drugs can be used to control HIV infection. However, caution is needed with SELY

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

vectors, as they are more sensitive to anti-HIV drugs like fusion inhibitors, reverse-transcriptase inhibitors, and integrase inhibitors, which should be avoided during transduction to ensure successful integration of SELY therapeutic DNAs into the cell's genome. During cell culture, any anti-HIV drugs disclosed in the art or approved by regulatory agencies (ex: FDA, EMA) are appropriate to control HIV infection.

Bone marrow engraftment

SELY CD34+ hematopoietic progenitor cells, whether from the patient's own body or from a donor, can sometimes fail to engraft sufficiently when administered through intravenous or intrabone injection. This issue can be tackled by preparing the patient with a conditioning regimen before the cell infusion. Any conditioning regimens—including low-intensity, low exposure, minimal intensity, very low-dose, single-dose, and those that result in short graft persistence, low level of chimerism, graft rejection, cancer relapse, or insignificant overall survival—disclosed in the relevant art literature and used before bone marrow transplantation are applicable here. Concerns like graft rejection, cancer relapse, or low survival rates, which are relevant in cancer treatments, are less worrisome for SELY cell engraftment in non-cancer patients. Conditioning regimens designed for toddlers and infants are suitable for adult patients as well.

When exposed to HIV, SELY cells generate SELY vectors capable of infecting CD4+ cells, leading to a rapid increase in SELY CD4+ cell numbers. Even a small number of SELY cells can trigger a chain reaction, significantly expanding the SELY cell population inside the patient. Long-term persistence of marrow grafts isn't a concern because SELY vectors can persist in peripheral CD4+ cells.

Shock-and-kill combination approach

The "shock-and-kill" strategy involves stimulating latent HIV-infected cells with latency-reversing agents, followed by controlling viral rebound with antiretroviral therapy and immune-mediated clearance of infected cells. To enhance this approach, patients can receive autologous or allogeneic SELY CD4+ cells or SELY vectors, followed by latency-reversing agent administration. Antiretroviral drugs are given only when HIV RNA levels surpass a certain ‘safety’ threshold (ex: 200,000 copies/mL). Upon HIV rebound, the virus infects CD4+ and SELY cells, prompting the production of more SELY vectors, which then infect nearby CD4+ cells. This competitive interaction between SELY vectors and HIV, along with the generation of HIV-resistant SELY CD4+ cells, helps suppress HIV infection.

Newly formed latent HIV reservoirs likely contain SELY therapeutic DNAs, converting them into latent therapeutic reservoirs, SELY vector-producing cells, or HIV-resistant cells. The newly produced SELY vectors, sharing the same envelope as HIV, are conditionally replicating and can act as an antigen-presenting platform. Furthermore, SELY vector-infected activated CD4+ T cells may become HIV-resistant, aiding in the generation of an effective anti-HIV immune response by presenting HIV antigens to CD8+ T cells and B cells [84], as CD4+ T cells play a crucial role in promoting B cells’ antibody production and priming cytotoxic and memory CD8+ T cells when combined with dendritic cells [85].

SELY Chimeric Antigen Receptor T cells

When CAR T cells are modified with SELY vectors, they become SELY CAR T cells. A certain percentage, P_{act} x 100 %, of these cells become HIV-resistant, while the remaining percentage, $(1-P_{act})$ x 100 %, act as a therapeutic reservoir. Transforming HIV-specific CD4+ CAR T cells [86–92] with the SELY system is highly appealing because when these cells encounter HIV-infected cells in various body areas, they get activated, undergo robust expansion, and start producing SELY vectors upon HIV infection (FIGURE 17).

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

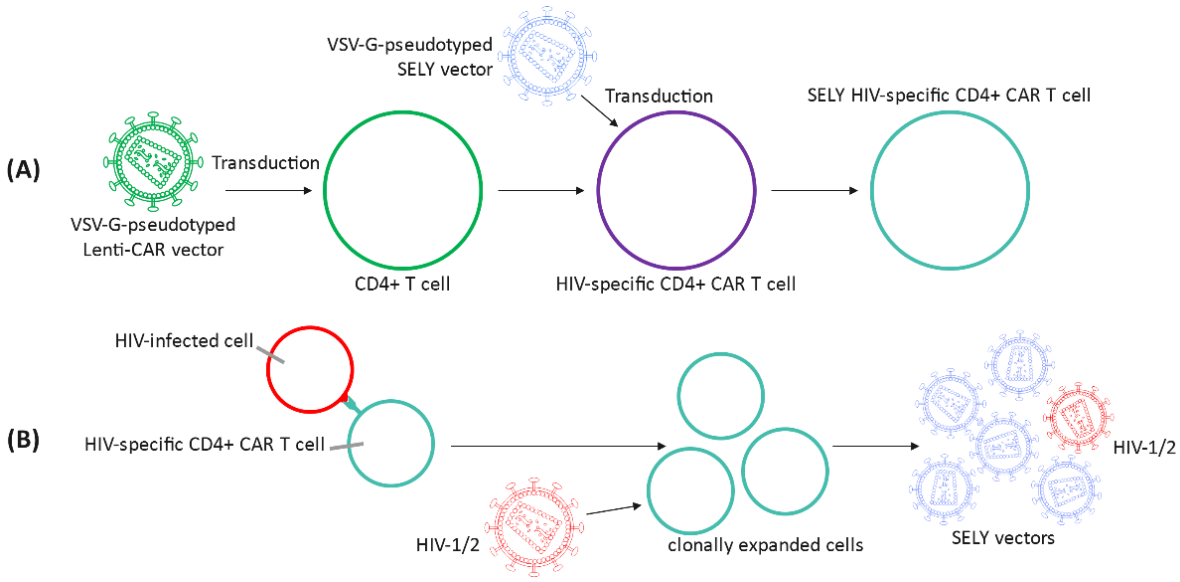

FIGURE 17: (A) illustrates the production of SELY HIV-specific CD4+ CAR T cells by modifying HIV-specific CD4+ CAR T cells with SELY vectors and a VSV-G-pseudotyped Lenti-CAR vector (the order of transduction is irrelevant). (B) Shows how SELY HIV-specific CD4+ CAR T cells rapidly expand after encountering an HIV-infected cell. Once infected by HIV, the expanded cells (part of the therapeutic reservoir subset) begin actively generating SELY vectors.

Chimeric Antigen Receptors (CARs) comprise an antigen binding domain, a hinge, a transmembrane domain, and cytoplasmic domain(s). CD4 extracellular domain [90], soluble CD4 [93], CD4 fragments [94], VH/VL segment derived from broadly neutralizing antibodies [91,92] are examples of suitable antigen binding domain for application here. Various cytoplasmic domains of chimeric antigen receptors described in existing literature are also suitable here. The cells don't necessarily need to have a CAR-encoding sequence integrated into their genome; transient expression of CAR after lipid-nanoparticle-mediated delivery of mRNAs encoding CAR is also effective.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest statement

The author declares no conflict of interest.

References

[1] D. Baltimore, “RNA-dependent DNA polymerase in virions of RNA tumour viruses,” *Nature* **226**, 1209–1211 (1970).

[2] H. M. Temin and S. Mizutani, “RNA-dependent DNA polymerase in virions of Rous sarcoma virus,” *Nature* **226**, 1211–1213 (1970).

[3] F. Barré-Sinoussi, J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vézinet-Brun, et al., “Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS),” *Science* **220**, 868–871 (1983).

[4] J. Coffin, A. Haase, J. A. Levy, L. Montagnier, S. Oroszlan, N. Teich, H. Temin, K. Toyoshima, H. Varmus, et al., “What to call the AIDS virus?,” *Nature* **321**, 10 (1986).

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

[5] C. van Schalkwyk, M. Mahy, L. F. Johnson, and J. W. Imai-Eaton, “Updated Data and Methods for the 2023 UNAIDS HIV Estimates,” *J Acquir Immune Defic Syndr* **95**, e1–e4 (2024).
[6] E. J. Arts and D. J. Hazuda, “HIV-1 antiretroviral drug therapy,” *Cold Spring Harb Perspect Med* **2**, a007161 (2012).
[7] R. M. Pinto, E. Hall, and R. Tomlin, “Injectable Long-Acting Cabotegravir-Rilpivirine Therapy for People Living With HIV/AIDS: Addressing Implementation Barriers From the Start,” *J Assoc Nurses AIDS Care* **34**, 216–220 (2023).
[8] L. Slama, R. Porcher, F. Linard, C. Chakvetadze, A. Cros, S. Carillon, L. Gallardo, J.-P. Viard, and J.-M. Molina, “Injectable long acting antiretroviral for HIV treatment and prevention: perspectives of potential users,” *BMC Infect Dis* **23**, 98 (2023).
[9] G. E. Gray, L.-G. Bekker, F. Laher, M. Malahleha, M. Allen, Z. Moodie, N. Grunenberg, Y. Huang, D. Grove, et al., “Vaccine Efficacy of ALVAC-HIV and Bivalent Subtype C gp120-MF59 in Adults,” *N Engl J Med* **384**, 1089–1100 (2021).
[10] B. F. Haynes, K. Wiehe, S. M. Alam, D. Weissman, and K. O. Saunders, “Progress with induction of HIV broadly neutralizing antibodies in the Duke Consortia for HIV/AIDS Vaccine Development,” *Curr Opin HIV AIDS* **18**, 300–308 (2023).
[11] B. Korber and W. Fischer, “T cell-based strategies for HIV-1 vaccines,” *Human Vaccines & Immunotherapeutics* **16**, 713–722 (2020).
[12] M. Giel-Moloney, M. Esteban, B. H. Oakes, M. Vaine, B. Asbach, R. Wagner, G. J. Mize, A. G. Spies, J. McElrath, et al., “Recombinant HIV-1 vaccine candidates based on replication-defective flavivirus vector,” *Sci Rep* **9**, 20005 (2019).
[13] E. Akbari, M. Seyedinkhorasani, and A. Bolhassani, “Conserved multiepitope vaccine constructs: A potent HIV-1 therapeutic vaccine in clinical trials,” *The Brazilian Journal of Infectious Diseases* **27**, 102774 (2023).
[14] G. Campos-Gonzalez, J. Martinez-Picado, T. Velasco-Hernandez, and M. Salgado, “Opportunities for CAR-T Cell Immunotherapy in HIV Cure,” *Viruses* **15**, 789 (2023).
[15] C. Sheridan, “Can gene editing drive out HIV and hepatitis viruses from inside cells?,” *Nat Biotechnol* **41**, 1491–1493 (2023).
[16] Y. Matsui and Y. Miura, “Advancements in Cell-Based Therapies for HIV Cure,” *Cells* **13**, 64 (2023).
[17] K. Chupradit, S. Moonmuang, S. Nangola, K. Kitidee, U. Yasamut, M. Mougel, and C. Tayapiwatana, “Current Peptide and Protein Candidates Challenging HIV Therapy beyond the Vaccine Era,” *Viruses* **9**, 281 (2017).
[18] M. C. Morris, V. Robert-Hebmann, L. Chaloin, J. Mery, F. Heitz, C. Devaux, R. S. Goody, and G. Divita, “A new potent HIV-1 reverse transcriptase inhibitor. A synthetic peptide derived from the interface subunit domains,” *J Biol Chem* **274**, 24941–24946 (1999).
[19] F. G. Hermann, H. Martinius, M. Egelhofer, T. Giroglou, T. Tonn, S. D. Roth, R. Zahn, P. Schult-Dietrich, A. Alexandrov, et al., “Protein Scaffold and Expression Level Determine Antiviral Activity of Membrane-Anchored Antiviral Peptides,” *Human Gene Therapy* **20**, 325–336 (2009).
[20] I. O. Gleenberg, A. Herschhorn, and A. Hizi, “Inhibition of the activities of reverse transcriptase and integrase of human immunodeficiency virus type-1 by peptides derived from the homologous viral protein R (Vpr),” *J Mol Biol* **369**, 1230–1243 (2007).
[21] A. B. Powell, Y. Ren, M. Korom, D. Saunders, P. J. Hanley, H. Goldstein, D. F. Nixon, C. M. Bollard, R. M. Lynch, et al., “Engineered Antigen-Specific T Cells Secreting Broadly Neutralizing Antibodies: Combining Innate and Adaptive Immune Response against HIV,” *Molecular Therapy - Methods & Clinical Development* **19**, 78–88 (2020).
[22] Q. Han, J. A. Jones, N. I. Nicely, R. K. Reed, X. Shen, K. Mansouri, M. Louder, A. M. Trama, S. M. Alam, et al., “Difficult-to-neutralize global HIV-1 isolates are neutralized by antibodies targeting open envelope conformations,” *Nat Commun* **10**, 2898 (2019).

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

[23] K. A. Delviks-Frankenberry, C. R. Ojha, K. J. Hermann, W.-S. Hu, B. E. Torbett, and V. K. Pathak, “Potent dual block to HIV-1 infection using lentiviral vectors expressing fusion inhibitor peptide mC46- and Vif-resistant APOBEC3G,” *Molecular Therapy - Nucleic Acids* **33**, 794–809 (2023).
[24] M. Delville, F. Touzot, C. Couzin, I. Hmitou, L. Djerroudi, A. Ouedrani, F. Lefrère, C. Tuchman-Durand, C. Mollet, et al., “Safety of CD34+ Hematopoietic Stem Cells and CD4+ T Lymphocytes Transduced with LVsh5/C46 in HIV-1 Infected Patients with High-Risk Lymphoma,” *Mol Ther Methods Clin Dev* **13**, 303–309 (2019).
[25] J. Zeidan, A. A. Sharma, G. Lee, A. Raad, R. Fromentin, S. Fourati, K. Ghneim, G. P. Sanchez, C. Benne, et al., “Infusion of CCR5 Gene-Edited T Cells Allows Immune Reconstitution, HIV Reservoir Decay, and Long-Term Virological Control,” preprint, Immunology, 2021.
[26] J. Herskovitz, M. Hasan, M. Patel, W. R. Blomberg, J. D. Cohen, J. Machhi, F. Shahjin, R. L. Mosley, J. McMillan, et al., “CRISPR-Cas9 Mediated Exonic Disruption for HIV-1 Elimination,” *EBioMedicine* **73**, 103678 (2021).
[27] A. J. Atkins, A. G. Allen, W. Dampier, E. K. Haddad, M. R. Nonnemacher, and B. Wigdahl, “HIV-1 cure strategies: why CRISPR?,” *Expert Opinion on Biological Therapy* **21**, 781–793 (2021).
[28] C. Yin, T. Zhang, X. Qu, Y. Zhang, R. Putatunda, X. Xiao, F. Li, W. Xiao, H. Zhao, et al., “In Vivo Excision of HIV-1 Provirus by saCas9 and Multiplex Single-Guide RNAs in Animal Models,” *Mol Ther* **25**, 1168–1186 (2017).
[29] J. Ding, Y. Liu, and Y. Lai, “Knowledge From London and Berlin: Finding Threads to a Functional HIV Cure,” *Front. Immunol.* **12**, 688747 (2021).
[30] R. K. Gupta, D. Peppa, A. L. Hill, C. Gálvez, M. Salgado, M. Pace, L. E. McCoy, S. A. Griffith, J. Thornhill, et al., “Evidence for HIV-1 cure after CCR5Δ32/Δ32 allogeneic haemopoietic stem-cell transplantation 30 months post analytical treatment interruption: a case report,” *The Lancet HIV* **7**, e340–e347 (2020).
[31] D. Peterhoff, “New case of HIV cure: joined forces of haploidentical stem cells and HLA-mismatched cord blood,” *Sig Transduct Target Ther* **8**, 241 (2023).
[32] J. K. Dickter, A. Aribi, A. A. Cardoso, S. Gianella, K. Gendzekhadze, S. Li, Y. Feng, A. Chaillon, G. M. Laird, et al., “HIV-1 Remission after Allogeneic Hematopoietic-Cell Transplantation,” *N Engl J Med* **390**, 669–671 (2024).
[33] J. Verheyen, A. Thielen, N. Lübke, M. Dirks, M. Widera, U. Dittmer, L. Kordelas, M. Däumer, D. C. M. de Jong, et al., “Rapid Rebound of a Preexisting CXCR4-tropic Human Immunodeficiency Virus Variant After Allogeneic Transplantation With CCR5 Δ32 Homozygous Stem Cells,” *Clin Infect Dis* **68**, 684–687 (2019).
[34] E. J. Tanner, S.-Y. Jung, J. Glazier, C. Thompson, Y. Zhou, B. Martin, H.-I. Son, J. L. Riley, and L. S. Weinberger, “Discovery and Engineering of a Therapeutic Interfering Particle (TIP): a combination self-renewing antiviral,” preprint, Synthetic Biology, 2019.
[35] E. Klimatcheva, V. Planelles, S. L. Day, F. Fulreader, M. J. Renda, and J. Rosenblatt, “Defective Lentiviral Vectors Are Efficiently Trafficked by HIV-1 and Inhibit Its Replication,” *Molecular Therapy* **3**, 928–939 (2001).
[36] W. An and A. Telesnitsky, “Effects of varying sequence similarity on the frequency of repeat deletion during reverse transcription of a human immunodeficiency virus type 1 vector,” *J Virol* **76**, 7897–7902 (2002).
[37] K. A. Delviks-Frankenberry, D. Ackerman, N. D. Timberlake, M. Hamscher, O. A. Nikolaitchik, W.-S. Hu, B. E. Torbett, and V. K. Pathak, “Development of Lentiviral Vectors for HIV-1 Gene Therapy with Vif-Resistant APOBEC3G,” *Mol Ther Nucleic Acids* **18**, 1023–1038 (2019).
[38] N. Srinivasakumar, “RRE-deleting self-inactivating and self-activating HIV-1 vectors for improved safety,” *PeerJ* **1**, e84 (2013).

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

[39] A. Duvergé and M. Negroni, “Pseudotyping Lentiviral Vectors: When the Clothes Make the Virus,” *Viruses* **12**, 1311 (2020).
[40] X.-Y. Zhang, V. F. La Russa, and J. Reiser, “Transduction of Bone-Marrow-Derived Mesenchymal Stem Cells by Using Lentivirus Vectors Pseudotyped with Modified RD114 Envelope Glycoproteins,” *J Virol* **78**, 1219–1229 (2004).
[41] N. Seddiki, J. Zaunders, C. Phetsouphanh, V. Brezar, Y. Xu, H. M. McGuire, M. Bailey, K. McBride, W. Hey-Cunningham, et al., “CD73+ CD127high Long-Term Memory CD4 T Cells Are Highly Proliferative in Response to Recall Antigens and Are Early Targets in HIV-1 Infection,” *Int J Mol Sci* **22**, 912 (2021).
[42] R. Zufferey, J. E. Donello, D. Trono, and T. J. Hope, “Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element Enhances Expression of Transgenes Delivered by Retroviral Vectors,” *J Virol* **73**, 2886–2892 (1999).
[43] J. K. Choi, N. Hoang, A. M. Vilardi, P. Conrad, S. G. Emerson, and A. M. Gewirtz, “Hybrid HIV/MSCV LTR Enhances Transgene Expression of Lentiviral Vectors in Human CD34 $^{+}$ Hematopoietic Cells,” *STEM CELLS* **19**, 236–246 (2001).
[44] B. L. Levine, L. M. Humeau, J. Boyer, R.-R. MacGregor, T. Rebello, X. Lu, G. K. Binder, V. Slepushkin, F. Lemiale, et al., “Gene transfer in humans using a conditionally replicating lentiviral vector,” *Proc Natl Acad Sci U S A* **103**, 17372–17377 (2006).
[45] P. Tebas, D. Stein, G. Binder-Scholl, R. Mukherjee, T. Brady, T. Rebello, L. Humeau, M. Kalos, E. Papasavvas, et al., “Antiviral effects of autologous CD4 T cells genetically modified with a conditionally replicating lentiviral vector expressing long antisense to HIV,” *Blood* **121**, 1524–1533 (2013).
[46] A. L. Szymczak-Workman, K. M. Vignali, and D. A. A. Vignali, “Design and Construction of 2A Peptide-Linked Multicistronic Vectors,” *Cold Spring Harb Protoc* **2012**, pdb.ip067876 (2012).
[47] Z. Liu, O. Chen, J. B. J. Wall, M. Zheng, Y. Zhou, L. Wang, H. Ruth Vaseghi, L. Qian, and J. Liu, “Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector,” *Sci Rep* **7**, 2193 (2017).
[48] W. Tang, I. Ehrlich, S. B. E. Wolff, A.-M. Michalski, S. Wolfl, M. T. Hasan, A. Luthi, and R. Sprengel, “Faithful Expression of Multiple Proteins via 2A-Peptide Self-Processing: A Versatile and Reliable Method for Manipulating Brain Circuits,” *Journal of Neuroscience* **29**, 8621–8629 (2009).
[49] G. Geng, B. Liu, C. Chen, K. Wu, J. Liu, Y. Zhang, T. Pan, J. Li, Y. Yin, et al., “Development of an Attenuated Tat Protein as a Highly-effective Agent to Specifically Activate HIV-1 Latency,” *Molecular Therapy* **24**, 1528–1537 (2016).
[50] E. Van Gulck, M. Pardons, E. Nijs, N. Verheyen, K. Dockx, C. Van Den Eynde, E. Battivelli, J. Vega, E. Florence, et al., “A truncated HIV Tat demonstrates potent and specific latency reversal activity,” *Antimicrob Agents Chemother* **67**, M. A. Martinez, Ed., e00417-23 (2023).
[51] S. Ngoi, A. Chien, and C. Lee, “Exploiting Internal Ribosome Entry Sites in Gene Therapy Vector Design,” *CGT* **4**, 15–31 (2004).
[52] R. Marques, R. Lacerda, and L. Romão, “Internal Ribosome Entry Site (IRES)-Mediated Translation and Its Potential for Novel mRNA-Based Therapy Development,” *Biomedicines* **10**, 1865 (2022).
[53] K. C. Straathof, M. A. Pulè, P. Yotnda, G. Dotti, E. F. Vanin, M. K. Brenner, H. E. Heslop, D. M. Spencer, and C. M. Rooney, “An inducible caspase 9 safety switch for T-cell therapy,” *Blood* **105**, 4247–4254 (2005).
[54] R. Solofondrazaintsoanirina, “DESNASEs: a group of RNASE-based safety switches for Controlled Cell Elimination” ([object Object], 2024).
[55] B. Philip, E. Kokalaki, L. Mekkaoui, S. Thomas, K. Straathof, B. Flutter, V. Marin, T. Marafioti, R. Chakraverty, et al., “A highly compact epitope-based marker/suicide gene for easier and safer T-cell therapy,” *Blood* **124**, 1277–1287 (2014).

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

[56] C. N. Barese, A. E. Krouse, M. E. Metzger, C. A. King, C. Traversari, F. C. Marini, R. E. Donahue, and C. E. Dunbar, “Thymidine kinase suicide gene-mediated ganciclovir ablation of autologous gene-modified rhesus hematopoiesis,” *Mol Ther* **20**, 1932–1943 (2012).
[57] R. Akkina, A. Banerjea, J. Bai, J. Anderson, M.-J. Li, and J. Rossi, “siRNAs, ribozymes and RNA decoys in modeling stem cell-based gene therapy for HIV/AIDS,” *Anticancer Res* **23**, 1997–2005 (2003).
[58] T. Zhang, T. Cheng, L. Wei, Y. Cai, A. E. Yeo, J. Han, Y. A. Yuan, J. Zhang, and N. Xia, “Efficient inhibition of HIV-1 replication by an artificial polycistronic miRNA construct,” *Virol J* **9**, 118 (2012).
[59] X. Li, N. Lobo, C. Bauser, and M. Fraser, “The minimum internal and external sequence requirements for transposition of the eukaryotic transformation vector piggyBac,” *Mol Gen Genomics* **266**, 190–198 (2001).
[60] S. A. Narayanavari and Z. Izsvák, “Sleeping Beauty transposon vectors for therapeutic applications: advances and challenges,” *Cell Gene Therapy Insights* **3**, 131–158 (2017).
[61] C. Miskey, “The Frog Prince: a reconstructed transposon from Rana pipiens with high transpositional activity in vertebrate cells,” *Nucleic Acids Research* **31**, 6873–6881 (2003).
[62] S. Y. Hwang, Y. H. Lee, M. U. Kuk, J. W. Kim, S. Oh, and J. T. Park, “Improvement of Tol2 Transposon System Enabling Efficient Protein Production in CHO Cells,” *Biotechnol Bioproc E* **26**, 767–775 (2021).
[63] D. Balciunas, K. J. Wangensteen, A. Wilber, J. Bell, A. Geurts, S. Sivasubbu, X. Wang, P. B. Hackett, D. A. Largaespada, et al., “Harnessing a High Cargo-Capacity Transposon for Genetic Applications in Vertebrates,” *PLoS Genet* **2**, e169 (2006).
[64] G. Sutrave, N. Xu, T. C. Y. Tang, A. Dolnikov, B. Gloss, D. J. Gottlieb, K. P. Micklethwaite, and K. Gowrishankar, “Characterizing piggyBat—a transposase for genetic modification of T cells,” *Molecular Therapy - Methods & Clinical Development* **25**, 250–263 (2022).
[65] Solofondrazaintsoanirina, Rojopitiavana, “Petit Plasmid: A Small, Versatile Antibiotic-Free Multi-Application DNA Vector Driving Biotech Innovation” (Zenodo, 2023).
[66] R. Solofondrazaintsoanirina, “Novel bacterial strains, suitable for use in large-scale fermentation, for the maintenance and replication of Petit Plasmid” ([object Object], 2024).
[67] A. Moretti, M. Ponzo, C. A. Nicolette, I. Y. Tcherepanova, A. Biondi, and C. F. Magnani, “The Past, Present, and Future of Non-Viral CAR T Cells,” *Front. Immunol.* **13**, 867013 (2022).
[68] S. R. Bartz and M. Emerman, “Human immunodeficiency virus type 1 Tat induces apoptosis and increases sensitivity to apoptotic signals by up-regulating FLICE/caspase-8,” *J Virol* **73**, 1956–1963 (1999).
[69] J. G. Julias, D. Hash, and V. K. Pathak, “E- vectors: development of novel self-inactivating and self-activating retroviral vectors for safer gene therapy,” *J Virol* **69**, 6839–6846 (1995).
[70] K. E. Langley, M. R. Villarejo, A. V. Fowler, P. J. Zamenhof, and I. Zabin, “Molecular basis of beta-galactosidase alpha-complementation,” *Proc Natl Acad Sci U S A* **72**, 1254–1257 (1975).
[71] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith, “Enzymatic assembly of DNA molecules up to several hundred kilobases,” *Nat Methods* **6**, 343–345 (2009).
[72] C. Engler, R. Kandzia, and S. Marillonnet, “A one pot, one step, precision cloning method with high throughput capability,” *PLoS One* **3**, e3647 (2008).
[73] M. Raman and K. Martin, “One solution for cloning and mutagenesis: In-Fusion® HD Cloning Plus,” *Nat Methods* **11**, iii–v (2014).
[74] W. R. Beyer, M. Westphal, W. Ostertag, and D. von Laer, “Oncoretrovirus and Lentivirus Vectors Pseudotyped with Lymphocytic Choriomeningitis Virus Glycoprotein: Generation, Concentration, and Broad Host Range,” *J Virol* **76**, 1488–1495 (2002).

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

[75] G. D. Trobridge, R. A. Wu, M. Hansen, C. Ironside, K. L. Watts, P. Olsen, B. C. Beard, and H.-P. Kiem, “Cocal-pseudotyped lentiviral vectors resist inactivation by human serum and efficiently transduce primate hematopoietic repopulating cells,” *Mol Ther* **18**, 725–733 (2010).
[76] A. S. Cockrell, H. Ma, K. Fu, T. J. McCown, and T. Kafri, “A trans-lentiviral packaging cell line for high-titer conditional self-inactivating HIV-1 vectors,” *Molecular Therapy* **14**, 276–284 (2006).
[77] S. Broussau, N. Jabbour, G. Lachapelle, Y. Durocher, R. Tom, J. Transfiguracion, R. Gilbert, and B. Massie, “Inducible Packaging Cells for Large-scale Production of Lentiviral Vectors in Serum-free Suspension Culture,” *Molecular Therapy* **16**, 500–507 (2008).
[78] M. Qamar Saeed, N. Dufour, C. Bartholmae, U. Sieranska, M. Knopf, E. Thierry, S. Thierry, O. Delelis, N. Grandchamp, et al., “Comparison Between Several Integrase-defective Lentiviral Vectors Reveals Increased Integration of an HIV Vector Bearing a D167H Mutant,” *Molecular Therapy - Nucleic Acids* **3**, e213 (2014).
[79] D. Karpova, M. P. Rettig, and J. F. DiPersio, “Mobilized peripheral blood: an updated perspective,” *F1000Res* **8**, F1000 Faculty Rev-2125 (2019).
[80] A. E. Sloan, H. Fung, J. Reese, L. Rgers, C. Murphy, H. Lazarus, B. Dropulic, and S. Gerson, “CTNI-49. PHASE I STUDY OF MGMT-P140K TRANSFECTED HEMATOPOETIC PROGENITOR CELLS COMBINED WITH TMZ/O6BG DOSE ESCALATION FOR NEWLY DIAGNOSED, UNMETHYLATED GLIOBLASTOMA: TOLERANCE AND EVIDENCE OF SURVIVAL BENEFIT,” *Neuro-Oncology* **22**, ii53–ii53 (2020).
[81] W. S. Lewis, V. Cody, N. Galitsky, J. R. Luft, W. Pangborn, S. K. Chunduru, H. T. Spencer, J. R. Appleman, and R. L. Blakley, “Methotrexate-resistant variants of human dihydrofolate reductase with substitutions of leucine 22. Kinetics, crystallography, and potential as selectable markers,” *J Biol Chem* **270**, 5057–5064 (1995).
[82] K. Hacke, J. A. Treger, B. T. Bogan, R. H. Schiestl, and N. Kasahara, “Genetic modification of mouse bone marrow by lentiviral vector-mediated delivery of hypoxanthine-Guanine phosphoribosyltransferase short hairpin RNA confers chemoprotection against 6-thioguanine cytotoxicity,” *Transplant Proc* **45**, 2040–2044 (2013).
[83] R. T. Mitsuyasu, J. A. Zack, J. L. Macpherson, and G. P. Symonds, “Phase I/II Clinical Trials Using Gene-Modified Adult Hematopoietic Stem Cells for HIV: Lessons Learnt,” *Stem Cells International* **2011**, 1–8 (2011).
[84] A. Jain, G. E. Canepa, M.-L. Liou, E. L. Fledderman, A. I. Chapoval, L. Xiao, I. Mukherjee, J. A. Galvin, P. N. Kumar, et al., “Multiple Treatment Interruptions and Protecting HIV-Specific CD4 T-Cells Enables Durable CTL Response and Viral Control,” preprint, HIV/AIDS, 2023.
[85] S. L. Swain, K. K. McKinstry, and T. M. Strutt, “Expanding roles for CD4+ T cells in immunity to viruses,” *Nat Rev Immunol* **12**, 136–148 (2012).
[86] R. E. Walker, C. M. Bechtel, V. Natarajan, M. Baseler, K. M. Hege, J. A. Metcalf, R. Stevens, A. Hazen, R. M. Blaese, et al., “Long-term in vivo survival of receptor-modified syngeneic T cells in patients with human immunodeficiency virus infection,” *Blood* **96**, 467–474 (2000).
[87] M. R. Roberts, L. Qin, D. Zhang, D. H. Smith, A. C. Tran, T. J. Dull, J. E. Groopman, D. J. Capon, R. A. Byrn, et al., “Targeting of human immunodeficiency virus-infected cells by CD8+ T lymphocytes armed with universal T-cell receptors,” *Blood* **84**, 2878–2889 (1994).
[88] J. Scholler, T. L. Brady, G. Binder-Scholl, W.-T. Hwang, G. Plesa, K. M. Hege, A. N. Vogel, M. Kalos, J. L. Riley, et al., “Decade-long safety and function of retroviral-modified chimeric antigen receptor T cells,” *Sci Transl Med* **4**, 132ra53 (2012).
[89] B. Liu, W. Zhang, B. Xia, S. Jing, Y. Du, F. Zou, R. Li, L. Lu, S. Chen, et al., “Broadly neutralizing antibody-derived CAR T cells reduce viral reservoir in individuals infected with HIV-1,” *J Clin Invest* **131**, e150211 (2021).
[90] S. G. Deeks, B. Wagner, P. A. Anton, R. T. Mitsuyasu, D. T. Scadden, C. Huang, C. Macken, D. D. Richman, C. Christopherson, et al., “A Phase II Randomized Study of HIV-Specific T-Cell Gene

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

Therapy in Subjects with Undetectable Plasma Viremia on Combination Antiretroviral Therapy," *Molecular Therapy* **5**, 788–797 (2002).

[91] A. Ali, S. G. Kitchen, I. S. Y. Chen, H. L. Ng, J. A. Zack, and O. O. Yang, "HIV-1-Specific Chimeric Antigen Receptors Based on Broadly Neutralizing Antibodies," *J Virol* **90**, 6999–7006 (2016).

[92] M. Hale, T. Mesojednik, G. S. Romano Ibarra, J. Sahni, A. Bernard, K. Sommer, A. M. Scharenberg, D. J. Rawlings, and T. A. Wagner, "Engineering HIV-Resistant, Anti-HIV Chimeric Antigen Receptor T Cells," *Mol Ther* **25**, 570–579 (2017).

[93] H. Haim, Z. Si, N. Madani, L. Wang, J. R. Courter, A. Princiotto, A. Kassa, M. DeGrace, K. McGee-Estrada, et al., "Soluble CD4 and CD4-Mimetic Compounds Inhibit HIV-1 Infection by Induction of a Short-Lived Activated State," *PLoS Pathog* **5**, T. J. Hope, Ed., e1000360 (2009).

[94] B. Dey, C. S. Del Castillo, and E. A. Berger, "Neutralization of Human Immunodeficiency Virus Type 1 by sCD4-17b, a Single-Chain Chimeric Protein, Based on Sequential Interaction of gp120 with CD4 and Coreceptor," *J Virol* **77**, 2859–2865 (2003).

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

Supplementary materials

Table S1: HIV-1 translation start sequences (extracted from AF033819.3)

Gene	Sequence
Gag	GACTAGCGGAGGCTAGAAGGAGAGAGATGG
Vif	AAGAAAAGCAAAGATCATTAGGGATTATGG
Vpr	AGTGTTACGAAACTGACAGAGGATAGATGG
Tat	CGTTACTCGACAGAGGAGAGCAAGAAATGG
Rev	CATAACAAAAGCCTTAGGCATCTCCTATGG
Vpu	TATCAAAGCAGTAAGTAGTACATGTAATGC
Env	AATAGAAAGAGCAGAAGACAGTGGCAATGA
Nef	GGGCTTGGAAAGGATTTTGCTATAAGATGG

Table S2: Exemplary components of HIV-1 transfer genome and SELY therapeutic sequences.

Name	Description	Sequence
R_U5_psi	R, U5, SD1, Psi sequences.	GGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTA ACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCT TCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATC CCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCG AACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCGACGCA GGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGC GACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGA GAGAG
tGag	Truncated Gag sequence containing internal in-frame stop codon(s).	**ATG**GGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCG ATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAA TTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGT TAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGG GACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCAT TATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGA TAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAA CAAAAGTAAGA
tGag255	Truncated Gag sequence (first 255-nt) containing internal in-frame stop codon(s).	**ATG**GGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCG ATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAA TTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGT TAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGG GACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCAT TATATAATACAGTAGCAACCC
tGag40	Truncated Gag sequence (first 40-nt).	**ATG**GGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAG
cPPT/CTS	Central polypurine tract and termination sequences. PPT.	TTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAG AATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAA AACAAATTACAAAATTCAAAATTTTCGGGTTTATTA
Env_RRE	Portion of the Env sequence comprising the Rev-Response	GATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGT GAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGC ACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGC AGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAA

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

	Element. Splice acceptor 7 (SA7).	GCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAA TTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATT GAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCA GCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAAC AGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTG CTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTG GAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACA CAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAA AGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGG AATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATA ATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTT TCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAG ACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGA AGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTG AACGGATCTCGACGGTATC
Env_RRE_SA7$_{enh}$	Portion of the Env sequence comprising the Rev-Response Element. The strength of the 3' splice site comprising Splice acceptor 7 (SA7) is enhanced. **Branch point.**	GATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGT GAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGC ACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGC AGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAA GCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAA TTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATT GAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCA GCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAAC AGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTG CTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTG GAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACA CAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAA AGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGG AATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATA ATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCGTTCTTT CTATTGCTAACCGTGTTCGTCAAGGTTATTCTCCTCTTTCTTTTCAGAC CCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAA GAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGA ACGGATCTCGACGGTATC
fU3_R_U5	Full length U3 sequence, R, and U5	GGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCA CTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAA GAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACT TCCCTGATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCA CTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAA GATAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTG TGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTG GAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGC ATCCGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCTACAAGGGA CTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTG GGGAGTGGCGAGCCCTCAGATCCTGCATATAAGCAGCTGCTTTTTGC CTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTC TGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTG AGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTA GAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAG

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

ΔU3_R_U5	U3 with a 400-nt deletion, R, U5	GGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCA CTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAC GAAAATAAGATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACC AGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTT AAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGT CTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCA GTGTGGAAAATCTCTAGCAG
SA3_tat216$_{ex1}$	Splice acceptor 3 (SA3), and first 216-nt of Tat exon 1 (encoding the first 72-aa of Tat).	CAGAATTGGGTGTCGACATAGCAGAATAGGCGTTACTCGACAGAGG AGAGCAAGAAATGGAGCCAGTAGATCCTAGACTAGAGCCCTGGAAG CATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAA AAGTGTTGCTTTCATTGCCAAGTTTGTTTCATAACAAAAGCCTTAGGC ATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCATC AGAACAGTCAGACTCATCAAGCTTCTCTATCAAAGCAG
SA3_tat198$_{ex1}$	Splice acceptor 3 (SA3), and first 198-nt of Tat exon 1 (encoding the first 66-aa of Tat).	CAGAATTGGGTGTCGACATAGCAGAATAGGCGTTACTCGACAGAGG AGAGCAAGAAATGGAGCCAGTAGATCCTAGACTAGAGCCCTGGAAG CATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAA AAGTGTTGCTTTCATTGCCAAGTTTGTTTCATAACAAAAGCCTTAGGC ATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCATC AGAACAGTCAGACTCATCAA

Table S3: Promoter 1 sequences.

Name	Description	Sequence
Prom1_CMV	human cytomegalovirus immediate early enhancer + immediate early promoter	GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTA GTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATG GCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAAT GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAA TGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT ATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCC CGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGC AGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTG GCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCC AAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAAT CAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAA TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGCGCGTTTT GCCTGTACT
Prom1_RSV	Rous sarcoma virus enhancer + promoter	TGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACGATGAG TTAGCAACATGCCTTACAAGGAGAGAAAAAGCACCGTGCATGCCGATT GGTGGAAGTAAGGTGGTACGATCGTGCCTTATTAGGAAGGCAACAGAC GGGTCTGACATGGATTGGACGAACCACTGAATTGCCGCATTGCAGAGA TATTGTATTTAAGTGCCTAGCTCGATACATAAAC
Prom1_B2M	Human beta-2-microglobulin promoter	GAAACCCTGCAGGGAATTCCCAAGCTGTAGTTATAAACAGAAGTTCTCC TTCTGCTAGGTAGCATTCAAAGATCTTAATCTTCTGGGTTTCCGTTTTCTC GAATGAAAAATGCAGGTCCGAGCAGTTAACTGGCTGGGGCACCATTAG CAAGTCACTTAGCATCTCTGGGGCCAGTCTGCAAAGCGAGGGGGCAGC CTTAATGTGCCTCCAGCCTGAAGTCCTAGAATGAGCGCCCGGTGTCCCA AGCTGGGGCGCGCACCCCAGATCGGAGGGCGCCGATGTACAGACAGC AAACTCACCCAGTCTAGTGCATGCCTTCTTAAACATCACGAGACTCTAA GAAAAGGAAACTGAAAACGGGAAAGTCCCTCTCTCTAACCTGGCACTG

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

		CGTCGCTGGCTTGGAGACAGGTGACGGTCCCTGCGGGCCTTGTCCTGA TTGGCTGGGCACGCGTTTAATATAAGTGGAGGCGTCGCGCTGGCGGGC
Prom1_U3	HIV-1 U3 sequence	CTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGT GGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACC AGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCT AGTACCAGTTGAGCCAGATAAGATAGAAGAGGCCAATAAAGGAGAGA ACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGG AGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATC ACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGACATC GAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTG GCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCATATAAG CAGCTGCTTTTTGCCTGTACT

Table S4: Exemplary transposon flanking sequences.

Name	Description	Sequence
Tol2_5_prime	5’ flanking region of Tol2 transposon	CAGAGGTGTAAAGTACTTGAGTAATTTTACTTGATTACTGTACTTAAGTA TTATTTTTGGGGATTTTTACTTTACTTGAGTACAATTAAAAATCAATACTT TTACTTTTACTTAATTACATTTTTTTAGAAAAAAAAGTACTTTTTACTCCT TACAATTTTATTTACAGTCAAAAAGTACTTATTTTTTGGAGATCACTT
Tol2_3_prime	3’ flanking region of Tol2 transposon	TAATACTCAAGTACAATTTTAATGGAGTACTTTTTTACTTTTACTCAAGTA AGATTCTAGCCAGATACTTTTACTTTTAATTGAGTAAAATTTTCCCTAAG TACTTGTACTTTCACTTGAGTAAAATTTTTGAGTACTTTTTACACCTCTG
SB_5_prime	5’ flanking region of Sleeping Beauty transposon	CAGTTGAAGTCGGAAGTTTACATACACTTAAGTTGGAGTCATTAAAACT CGTTTTTCAACTACTCCACAAATTTCTTGTTAACAAACAATAGTTTTGGC AAGTCAGTTAGGACATCTACTTTGTGCATGACACAAGTCATTTTTCCAAC AATTGTTTACAGACAGATTATTTCACTTATAATTCACTGTATCACAATTCC AGTGGGTCAGAAGTTTACATACACTAAGT
SB_3_prime	3’ flanking region of Sleeping Beauty transposon	ATTGAGTGTATGTAAACTTCTGACCCACTGGGAATGTGATGAAAGAAAT AAAAGCTGAAATGAATCATTCTCTCTACTATTATTCTGATATTTCACATTC TTAAAATAAAGTGGTGATCCTAACTGACCTAAGACAGGGAATTTTTACT AGGATTAAATGTCAGGAATTGTGAAAAAGTGAGTTTAAATGTATTTGGC TAAGGTGTATGTAAACTTCCGACTTCAACTG
ePB_5_prime	5’ flanking region of enhanced piggyback transposon	CGCAGCTAGATTAACCCTAGAAAGATAGTCTGCGTAAAATTGACGCATG CATTCTTGAAATATTGCTCTCTCTTTCTAAATAGCGCGAATCCGTCGCTG TGCATTTAGGACATCTCAGTCGCCGCTTGGAGCTCCCGTGAGGCGTGCT TGTCAATGCGGTAAGTGTCACTGATTTTGAACTATAACAACCGCGTGAG TCAAAATGACGCATGATTATCTTTTACGTGACTTTTAAGATTTAACTCAT ACGATAATTATATTGTTATTTCGTGTTCTACTTACGTGATAACTTATTATA TATATATTTTCTTGTTATAGATATCCTT
ePB_3_prime	3’ flanking region of enhanced piggyback transposon	CGATAAAAGTTTTGTTACTTTATAGAAGAAATTTTGAGTTTTTGTTTTTTT TTAATAAATAAATAAACATAAATAAATTGTTTGTTGAATTTATTATTAGT ATGTAAGTGTAAATATAATAAAACTTAATATCTATTCAAATTAATAAATA AACCTCGATATACAGACCGATAAAACACATGCGTCAATTTTACGCATGA TTATCTTTAACGTACGTCACAATATGATTATCTTTCTAGGGTTAATCTAG CTGC

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

Table S5: Polyadenylation sequences.

Name	Description	Sequence
L_SV40_polyA	Late SV40 polyadenylation sequence	CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAAT GCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTAT TTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCAT TCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGC AAGTAAAACCTCTACAAATGTGGTA
BGH_polyA	Bovine Growth Hormone polyadenylation sequence	CTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTT CCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAG GAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTG GGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAG GCATGCTGGGGATGCGGTGGGCTCTATGG

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana

In Memory of

RASOLOFONIRINA Marcel (Sely)

July 23, 1957 – February 13, 2024

My Dad was an exceptional father, devoted, caring, and loving with all his heart. He worked tirelessly to uplift his family into the middle class, overcoming myriad difficulties and life's most bitter challenges, one after another. Despite immense sacrifices, he always prevailed. I owe everything I am to you, Dad.

Copyright © 2024 Solofondrazaintsoanirina Rojoptiavana